An Introduction to Configuration Interaction Theory

C. David Sherrill
School of Chemistry and Biochemistry
Georgia Institute of Technology

1995

Contents

1   Introduction and Notation

These notes attempt to present the essential ideas of configuration interaction (CI) theory in a fairly detailed mathematical framework. Of all the *ab initio* methods, CI is probably the easiest to understand—and perhaps one of the hardest to implement efficiently on a computer! The next two sections explain what the CI method is: the matrix formulation of the Schrödinger equation $\hat{H}\Psi = E\Psi$. The remaining sections describe various simplifications, approximations, and computational techniquies.

I have attempted to use a uniform notation throughout these notes. Much of the notation is consistent with that of Szabo and Ostlund, *Modern Quantum Chemistry* [1]. Below are listed several of the commonly-used symbols and their meanings.

N The number of electrons in the system.

n_α The number of alpha electrons.

n_β The number of beta electrons.

n The number of orbitals in the one-particle basis set.

δ_{ij} Kronecker delta function, equal to one if $i = j$ and zero otherwise.

$\hat{H}$ The exact nonrelativistic Hamiltonian operator.

H The Hamiltonian matrix, i.e. the matrix form of $\hat{H}$, in whatever N-electron basis is currently being used.

H_{ij} The i, j-th element of **H**, equal to $\langle \Phi_i | \hat{H} | \Phi_j \rangle$, where Φ_i and Φ_j are N-electron CI basis functions.

$\mathbf{x}_i$ The space and spin coordinates of particle i.

$\mathbf{r}_i$ The spatial coordinates of particle i.

ϕ_i The i-th one-particle basis function (orbital). Usually denotes a spin-orbital obtained from a Hartree-Fock procedure. May also be written simply as i.

χ_i The i-th one-particle basis function (orbital). Usually denotes an atomic spin-orbital.

$|\Phi_i\rangle$ The i-th N-electron basis function. Usually denotes a single Slater determinant, but may also be a configuration state function (CSF).

$|\Psi\rangle$ Usually denotes an eigenfunction of **H**. The exact nonrelativistic wavefunction if a complete basis is used in the expansion of $\hat{H}$.

$|\Phi_a^r\rangle$ An N-electron basis function which differs from some reference function $|\Phi_0\rangle$ by the replacement of spin-orbital a by spin-orbital r. Usually implies a single Slater determinant.

$|ab\ldots c\rangle$ A Slater determinant with spin-orbitals a, b, $\ldots c$ occupied, i.e.

$$|\phi_a\phi_b\ldots\phi_c\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_a(\mathbf{x}_1) & \phi_b(\mathbf{x}_1) & \ldots & \phi_c(\mathbf{x}_1) \\ \phi_a(\mathbf{x}_2) & \phi_b(\mathbf{x}_2) & \ldots & \phi_c(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \phi_a(\mathbf{x}_N) & \phi_b(\mathbf{x}_N) & \ldots & \phi_c(\mathbf{x}_N) \end{vmatrix} \tag{1.1}$$

$\langle i|\hat{h}|j\rangle$ One-electron integral in physicists' notation (i and j are spin-orbitals). More explicitly, this is

$$\langle i|\hat{h}|j\rangle = \int \phi_i^*(\mathbf{x}_1)\hat{h}(\mathbf{x}_1)\phi_j(\mathbf{x}_1)d\mathbf{x}_1 \tag{1.2}$$

$[i|\hat{h}|j]$ One-electron integral in chemists' notation, where i and j are spin-orbitals. Equivalent to $\langle i|\hat{h}|j\rangle$.

$(i|\hat{h}|j)$ One-electron integral in chemists' notation (i and j are *spatial* orbitals).

$\langle ij||kl\rangle$ Antisymmetrized two-electron integral, equal to $\langle ij|kl\rangle - \langle ij|lk\rangle$.

$\langle ij|kl \rangle$ A simple two-electron integral, in physicists' notation, where i, j, k, and l are spin-orbitals. This is

$$\langle ij|kl \rangle = \int \phi_i^*(\mathbf{x}_1)\phi_j^*(\mathbf{x}_2)\frac{1}{\mathbf{r}_{12}}\phi_k(\mathbf{x}_1)\phi_l(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 \qquad (1.3)$$

$[\,ij|kl\,]$ A simple two-electron integral in chemists' notation, where i, j, k, and l are spin-orbitals. This is

$$[ij|kl] = \int \phi_i^*(\mathbf{x}_1)\phi_j(\mathbf{x}_1)\frac{1}{\mathbf{r}_{12}}\phi_k^*(\mathbf{x}_2)\phi_l(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 \qquad (1.4)$$

$(ij|kl)$ A simple two-electron in chemists' notation where i, j, k, and l are spatial orbitals. This is

$$(ij|kl) = \int \phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\frac{1}{\mathbf{r}_{12}}\phi_k^*(\mathbf{r}_2)\phi_l(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \qquad (1.5)$$

$a_i^\dagger$ Second-quantized creation operator for orbital i.

a_i Second-quantized annihilation operator for orbital i.

2 Fundamental Concepts

2.1 Scope of the Method

Configuration interaction (CI) is a method for solving the nonrelativistic Schrödinger equation

$$\hat{H}\Psi(\mathbf{r},\mathbf{R}) = \left\{\sum_A \frac{1}{2M_A}\nabla_A^2 + \sum_i \frac{1}{2}\nabla_i^2 + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{Ai}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}}\right\}\Psi(\mathbf{r},\mathbf{R}) = E\Psi(\mathbf{r},\mathbf{R})$$

$$(2.6)$$

where i, j denote electrons and A, B denote nuclei, with $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $R_{Ai} = |\mathbf{R}_A - \mathbf{r}_i|$, and $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$. Typical applications of the CI method employ the Born-Oppenheimer approximation, whereby the the motions of the electrons are treated as uncoupled from those of the nuclei. Thus the "electronic" Shrödinger equation is solved at discrete sets of *fixed* nuclear positions

$$\hat{H}_e\Psi_e(\mathbf{r};\mathbf{R}) = \left\{-\frac{1}{2}\sum_i\nabla_i^2 - \sum_{A,i}\frac{Z_A}{R_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}}\right\}\Psi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R})\Psi_e(\mathbf{r};\mathbf{R}) \quad (2.7)$$

The Born-Oppenheimer approximation is invoked so often in computational quantum chemistry that the subscripts in the preceeding equation are usually suppressed and the equation is written simply as $\hat{H}\Psi = E\Psi$. However, it is important to remember that the electronic energy E_e is an artifact of the Born-Oppenheimer approximation and is not as physically meaningful as the *total energy* of a system. Within the Born-Oppenheimer approximation, we estimate the total energy by adding the nuclear-nuclear repulsion energy and the nuclear kinetic energy to the total electronic energy E_e of equation (2.7).

While the CI method can be extended to incorporate some relativistic effects (e.g. spin-orbit terms), this is not generally done; these notes will be concerned only with the nonrelativistic Hamiltonian (2.7).

2.2 Why Configuration Interaction?

In the first paper on quantum mechanics, Heisenberg used matrix mechanics to calculate the frequencies and intensities of spectral lines [2]. Later, when Schrödinger discovered wave mechanics, it was quickly shown that the Schrödinger and Heisenberg approaches are mathematically equivalent [3, 4]. Given the ease with which matrices may be implemented on a computer, it is entirely natural to attempt to solve the molecular time-independent Schrödinger equation $\hat{H}\Psi = E\Psi$ using matrix mechanics.

Matrix mechanics requires that we choose a vector space for the expansion of the problem. For the case of an N-electron molecule, our wavefunction must be expanded in a basis of N-particle functions (the nuclei need not be considered in the electronic wavefunction, if we have invoked the Born-Oppenheimer approximation). How do we construct the N-particle basis functions? Here we follow the arguments of Szabo and Ostlund [1], p. 60. Assume we have a complete set of functions $\{\chi_i(x_1)\}$ of a single variable x_1. Then any arbitrary function of that variable can be expanded exactly as

$$\Phi(x_1) = \sum_i a_i \chi_i(x_1). \tag{2.1}$$

How can we expand a function of *two* variables x_1 and x_2 which have the same domain? If we hold x_2 fixed, then

$$\Phi(x_1, x_2) = \sum_i a_i(x_2)\chi_i(x_1). \tag{2.2}$$

Now note that each expansion coefficient $a_i(x_2)$ is a function of a single variable, which can be expanded as

$$a_i(x_2) = \sum_j b_{ij}\chi_j(x_2). \tag{2.3}$$

Substituting this expression into the one for $\Phi(x_1, x_2)$, we now have

$$\Phi(x_1, x_2) = \sum_{ij} b_{ij}\chi_i(x_1)\chi_j(x_2) \tag{2.4}$$

a process which can obviously be extended for $\Phi(x_1, x_2, \ldots, x_N)$.

Let us now collect the spin and space coordinates of an electron into a variable $\mathbf{x}$. We can write a spin orbital as $\chi(\mathbf{x})$. The result analogous to equation (2.4) for a system of N electrons is

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \sum_{ij\ldots N} b_{ij\ldots N} \chi_i(\mathbf{x}_1) \chi_j(\mathbf{x}_2) \ldots \chi_N(\mathbf{x}_N) \tag{2.5}$$

However, the wavefunction must be antisymmetric with respect to the exchange of the coordinates of any two electrons[1] For the two-particle case, the requirement

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = -\Phi(\mathbf{x}_2, \mathbf{x}_1) \tag{2.6}$$

implies that $b_{ij} = -b_{ji}$ and $b_{ii} = 0$, or

$$\begin{aligned} \Phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{j>i} b_{ij}[\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)] \\ &= \sum_{j>i} 2^{1/2} b_{ij} |\chi_i \chi_j\rangle \end{aligned} \tag{2.7}$$

More generally, an arbitrary N-electron wavefunction can be expressed exactly as a linear combination of all possible N-electron Slater determinants formed from a complete set of spin orbitals $\{\chi_i(\mathbf{x})\}$. If we solve the matrix mechanics problem $\mathbf{H}|\Psi\rangle = E|\Psi\rangle$ in a complete basis of N-electron functions as just described, we will obtain all electronic eigenstates of the system exactly. If our N-electron basis functions are denoted $|\Phi_i\rangle$, the eigenvectors of $\mathbf{H}$ are given as

$$|\Psi_j\rangle = \sum_{i}^{I} c_{ij} |\Phi_i\rangle \tag{2.8}$$

if there are I possible N-electron basis functions (I will be infinite if we actually have a complete set of one electron functions χ_i). The matrix $\mathbf{H}$ is constructed so that $H_{ij} = \langle \Phi_i | H | \Phi_j \rangle$ for $i, j = 1, 2, \ldots, I$. The matrix elements H_{ij} may be written in terms of one- and two-electron integrals according to "Slater's rules," as discussed in section 2.4.

---

[1] According to the antisymmetry principle for fermions, of which the Pauli principle is a direct result.

The N-electron basis functions $|\Phi_i\rangle$ can be written as substitutions or "excitations" from the Hartree-Fock "reference" determinant, i.e.

$$|\Psi\rangle = c_0|\Phi_0\rangle + \sum_{ra} c_a^r|\Phi_a^r\rangle + \sum_{a<b,r<s} c_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \sum_{r<s<t,a<b<c} c_{abc}^{rst}|\Phi_{abc}^{rst}\rangle + \ldots \qquad (2.9)$$

where $|\Phi_a^r\rangle$ means the Slater determinant formed by replacing spin-orbital a in $|\Phi_0\rangle$ with spin orbital r, etc. Every N-electron Slater determinant can be described by the set of N spin orbitals from which it is formed, and this set of orbital occupancies is often referred to as a "configuration." Thus the "configuration interaction" method is, in its most straigtforward implementation, nothing more or less than the matrix mechanics solution of the time-independent non-relativistic electronic Schrödinger equation $\hat{H}\Psi = E\Psi$. One of the great strengths of the CI method is its generality; the formalism applies to excited states, to open-shell systems, and to systems far from their equilibrium geometries. By contrast, traditional single-reference perturbation theory and coupled-cluster approaches generally *assume* that the reference configuration is dominant, and they may fail when it is not.

In practice, one does not have a complete set of one-particle basis functions $\{\chi_i(\mathbf{x})\}$; typically one assumes that the incomplete one-electron basis set is large enough to give useful results, and the CI procedure is not modified. The quality of the one-particle basis set can be checked by comparing the results of calculations using progressively larger basis sets.

It is also possible to reduce the size of the N-electron basis set. If we desire only wavefunctions of a given spin and/or spatial symmetry, as is usually the case, we need include only those N-electron basis functions of that symmetry, since the Hamiltonian matrix is block-diagonal according to space and spin symmetries. This point is discussed further in section 4.1. If one performs the matrix mechanics calculation using a given set of one-particle functions $\{\chi_i(\mathbf{x})\}$ and all possible N-electron basis functions $\{|\Phi_i\rangle\}$ (possibly symmetry-restricted), the procedure is called "full CI." The full CI corresponds to solving Schrödinger's equation exactly *within* the space spanned by the specified one-electron basis. If the one-electron basis is complete (it never is in practice, but it may be in theory), then the

procedure is called a "complete CI" [5].

Unfortunately, even with an incomplete one-electron basis, a full CI is computationally intractable for any but the smallest systems, due to the vast number of N-electron basis functions required (the size of the CI space is discussed in section 4.4). The CI space must be reduced somehow—hopefully in such a way that the approximate CI wavefunction and energy are as close as possible to the exact values. The effective reduction of the CI space is a major concern in CI theory, and we will discuss some of the more popular approaches in these notes.

By far the most common CI approximation is the truncation of the CI space expansion according to excitation level relative to the reference state (equation 2.9). The widely-employed CI singles and doubles (CISD) wavefunction includes only those N-electron basis functions which represent single or double excitations relative to the reference state. Since the Hamiltonian operator includes only one- and two-electron terms, only singly and doubly excited configurations can interact directly with the reference, and they typically account for about 95% of the correlation energy in small molecules at their equilibrium geometries [6]. Truncation of the CI space according to excitation class is discussed more thoroughly in section 4.2.

2.3 The Correlation Energy

Approximate CI methods can be judged according to what fraction of the correlation energy they recover. The correlation energy is defined as the difference between the energy in the Hartree-Fock limit (E_{HF}) and the exact nonrelativistic energy of a system ($\mathcal{E}_0$)

$$E_{\text{corr}} = \mathcal{E}_0 - E_{HF} \tag{2.10}$$

This energy will always be negative because the Hartree-Fock energy is an upper bound to the exact energy (this is guaranteed by the variational theorem, as explained in section 3). The exact nonrelativistic energy $\mathcal{E}_0$ could, in principle, be calculated by performing a full CI in a complete one-electron basis set. If we have an incomplete one-electron basis set, then we can only compute the *basis set*

correlation energy, which is the correlation energy for a given one-electron basis. For convenience, the basis set correlation energy is often simply referred to as the correlation energy.

The correlation energy is the energy recovered by fully allowing the electrons to avoid each other; Hartree-Fock improperly treats interelectron repulsions in an averaged way.[2] However, there is some inconsistency in this line of thinking. When a molecule is pulled apart, the electrons shouldn't need to avoid each other as much, so the magnitude of the correlation energy should decrease. In fact, the opposite is true, as shown by the basis set correlation energies given in Table 1 for H_2O at three different geometries.

Table 1: Electron Correlation in H_2O with a DZ Basis.

Geometry	E_{corr} (hartree)[a]
R_e	-0.148028
1.5 R_e	-0.210992
2.0 R_e	-0.310067

[a]Data from reference [6].

The correlation energy increases at stretched geometries, because our definition of the correlation energy in equation (2.10) includes not only the concept of electrons avoiding each other, which is called the "dynamical" correlation energy, but also a more subtle effect called the "nondynamical," or "static" correlation energy. Nondynamical correlation energy reflects the inadequacy of a single reference in describing a given molecular state, and is due to nearly degenerate states or rearrangement of electrons within partially filled shells. Shavitt [7] has pointed out this deficiency in the definition of the correlation energy, and has suggested that perhaps a multiconfigurational Hartree-Fock method may be more useful in the definition of correlation energy.

---

[2]Some texts talk about the "average" electron repulsion term in the Fock operator; I find this misleading in that Hartree-Fock uses the *same* instantaneous interelectron repulsion term as CI—it's the same Hamiltonian! The restriction to a single Slater determinant is what causes the averaging of interelectron repulsions.

Siegbahn [8] offers the following explanation of the difference between dynamical and nondynamical correlation energies:

> "In many situations it is further convenient to subdivide the correlation energy into two parts with different physical origins. For chemical reactions where bonds are broken and formed, and for most excited states, the major part of the correlation energy can be obtained by adding only a few extra configurations besides the Hartree-Fock configuration. This part of the correlation energy is due to near degeneracy between different configurations and has its origin quite often in artifacts of the Hartree-Fock approximation. The physical origin of the second part of the correlation energy is the dynamical correlation of the motion of the electrons and is therefore sometimes called the dynamical correlation energy. Since the Hamiltonian operator contains only one- and two-particle operators this part of the correlation energy can be very well described by single and double replacements from the leading, near degenerate, reference configurations."

2.4   Slater's Rules

Whether we perform a full CI or only a limited CI, we must be able to express $\hat{H}$ in matrix form so that we can diagonalize it and obtain the eigenvectors and eigenvalues of interest. In this section we discuss Slater's rules (or the Slater-Condon rules [9, ?, 10]), which allow us to express matrix elements $H_{ij} = \langle \Phi_i | \hat{H} | \Phi_j \rangle$ in terms of one- and two-electron integrals. At the moment, we will express these results in terms spin-orbitals using physicist's notation. The one-electron integrals are written as

$$\langle i | \hat{h} | j \rangle = \int \phi_i^*(\mathbf{r}_1) \hat{h}(\mathbf{r}_1) \phi_j(\mathbf{r}_1) d\mathbf{r}_1 \tag{2.11}$$

and the two-electron integrals are written as

$$\langle ij || kl \rangle = \langle ij | kl \rangle - \langle ij | lk \rangle \tag{2.12}$$

where

$$\langle ij|kl\rangle = \int \phi_i^*(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\frac{1}{\mathbf{r}_{12}}\phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.13}$$

Before Slater's rules can be used, the two Slater determinants must be arranged in *maximum coincidence.* Remember that switching columns in a determinant introduces a minus sign. For instance, to calculate $\langle\Phi_1|\hat{H}|\Phi_2\rangle$, where we have

$$|\Phi_1\rangle = |abcd\rangle \tag{2.14}$$

$$|\Phi_2\rangle = |crds\rangle \tag{2.15}$$

then we must *first* interchange columns of $|\Phi_1\rangle$ or $|\Phi_2\rangle$ to make the two determinants look as much alike as possible. For example, we may rearrange $|\Phi_2\rangle$ as

$$|\Phi_2\rangle = |crds\rangle = -|crsd\rangle = |srcd\rangle \tag{2.16}$$

After the determinants are in maximum coincidence, we see how many spin orbitals they differ by, and we then use the following rules:

1. Identical Determinants: If the determinants are identical, then

$$\langle\Phi_1|\hat{H}|\Phi_1\rangle = \sum_m^N\langle m|\hat{h}|m\rangle + \sum_{m>n}^N\langle mn||mn\rangle \tag{2.17}$$

2. Determinants that Differ by One Spin Orbital:

$$\begin{aligned}
|\Phi_1\rangle &= |\cdots mn\cdots\rangle \\
|\Phi_2\rangle &= |\cdots pn\cdots\rangle \\
\langle\Phi_1|\hat{H}|\Phi_2\rangle &= \langle m|\hat{h}|p\rangle + \sum_n^N\langle mn||pn\rangle
\end{aligned} \tag{2.18}$$

3. Determinants that Differ by Two Spin Orbitals:

$$\begin{aligned}
|\Phi_1\rangle &= |\cdots mn\cdots\rangle \\
|\Phi_2\rangle &= |\cdots pq\cdots\rangle \\
\langle\Phi_1|\hat{H}|\Phi_2\rangle &= \langle mn||pq\rangle
\end{aligned} \tag{2.19}$$

4. Determinants that differ by More than Two Spin Orbitals:

$$|\Phi_1\rangle = |\cdots mno\cdots\rangle \qquad (2.20)$$
$$|\Phi_2\rangle = |\cdots pqr\cdots\rangle$$
$$\langle\Phi_1|\hat{H}|\Phi_2\rangle = 0$$

The derivation of these rules can be found in Szabo and Ostlund [1], section 2.3.4 (pp. 74-81).

3 The Variational Theorem

3.1 The Method of Linear Variations

In this section we show that the method of linear variations (also called the Ritz method [11]) is equivalent to the matrix formulation $\mathbf{Hc} = E\mathbf{c}$ of the Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$. Our treatment is similar to that of Szabo and Ostlund [1], p. 116. The linear variation method states that, given the linear expansion

$$|\Psi\rangle = \sum_i c_i |\Phi_i\rangle \tag{3.1}$$

we vary the coefficients c_i so that we minimize $E = \langle\Psi|\hat{H}|\Psi\rangle / \langle\Psi|\Psi\rangle$. We begin by requiring that the wavefunction be normalized so that $\langle\Psi|\Psi\rangle = 1$. Normalization means that we cannot minimize E simply by solving

$$\frac{\delta}{\delta c_k}\langle\Psi|\hat{H}|\Psi\rangle = 0 \qquad k = 1, 2, \ldots, N \tag{3.2}$$

because the c_i's are not independent. In this case we have a constrained minimization, so we apply Lagrange's method of undetermined multipliers, and we minimize the functional

$$\mathcal{L} = \langle\Psi|\hat{H}|\Psi\rangle - E(\langle\Psi|\Psi\rangle - 1) \tag{3.3}$$

which has the same minimum as E when $|\Psi\rangle$ is normalized. When we substitute equation (3.1) into equation (3.3), we obtain

$$\mathcal{L} = \sum_{ij} c_i^* c_j \langle\Phi_i|\hat{H}|\Phi_j\rangle - E\left(\sum_{ij} c_i^* c_j \langle\Phi_i|\Phi_j\rangle - 1\right) \tag{3.4}$$

which we may rewrite as

$$\mathcal{L} = \sum_{ij} c_i^* c_j H_{ij} - E\left(\sum_{ij} c_i^* c_j S_{ij} - 1\right) \tag{3.5}$$

where of course $H_{ij} = \langle\Phi_i|\hat{H}|\Phi_j\rangle$ and $S_{ij} = \langle\Phi_i|\Phi_j\rangle$. Now set the first variation in $\mathcal{L}$ equal to zero:

$$\delta\mathcal{L} = \sum_{ij} \delta c_i^* c_j H_{ij} - E\sum_{ij} \delta c_i^* c_j S_{ij} + \sum_{ij} c_i^* \delta c_j H_{ij} - E\sum_{ij} c_i^* \delta c_j S_{ij} = 0 \tag{3.6}$$

Since the summations run over all i and j, and since $H_{ij} = H_{ji}^*$ and $S_{ij} = S_{ji}^*$, we can simplify to

$$\delta\mathcal{L} = \sum_i \delta c_i^* \left[\sum_j H_{ij} c_j - E S_{ij} c_j \right] + \text{complex conj.} = 0 \qquad (3.7)$$

Since each term is a sum of a number and its complex conjugate, the imaginary parts will cancel. However, the real part will not necessarily be zero; in fact, since all the δc_i's are arbitrary (that is the whole point of using Lagrange's method), then for $\delta\mathcal{L}$ to be zero, the term in brackets must be zero. We may rewrite this condition as a matrix equation

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c} \qquad (3.8)$$

If the basis functions $\{\, |\Phi_i\rangle \,\}$ are chosen orthonormal (as is usually the case), then $\mathbf{S} = \mathbf{I}$ the identity matrix, and we have $\mathbf{H}\mathbf{c} = E\mathbf{c}$. Of course $\mathbf{c}$ is the column-vector representation of $|\Psi\rangle$ in the basis $\{\, |\Phi_i\rangle \,\}$. We thus have two equivalent ways of viewing a CI—either as the matrix formulation of the Schrödinger equation within the given linear vector space of N-electron basis functions, *or* as the minimization of the energy with respect to the linear expansion coefficients c_i of (3.1), subject to the constraint that the wavefunction remain normalized. Another way of viewing the results of this section is to note that only eigenvectors of the Hamiltonian matrix $\mathbf{H}$ are stable with respect to variations in the linear expansion coefficients.

At this point it is reasonable to ask *why* we wish to minimize the energy by varying the coefficients in equation (3.1). How do we know that this will give us the best estimate of the wavefunction? There are two answers to this. First, as we have just shown, minimizing the energy by variation of the linear expansion coefficients gives the Schrödinger equation in matrix form; thus the procedure is justified *a posteriori* by the validity of its result. The other reason is that, for the ground state, the linear expansion in equation (3.1) gives an expectation value for the energy E which is always an *upper bound* to the exact nonrelativistic ground state energy $\mathcal{E}_0$. We will prove this assertion in the next section; the result is called the Variational Theorem. The best estimate of E, then, is the minimum value which can be obtained by varying the coefficients in equation (3.1) (while

also maintaining normalization). These arguments also hold for excited states, so long as each excited state is made orthogonal to all lower states.

3.2   Variational Theorem for the Ground State

One particularly nice feature of the CI method is that the calculated lowest energy eigenvalue is always an upper bound to the exact ground state energy. Our approximate wavefunction $|\Theta\rangle$ can always be expressed as a linear combination of the exact nonrelativistic eigenvectors $|\Psi_i\rangle$, which span the entire N-electron space.

$$|\Theta\rangle = \sum_i c_i |\Psi_i\rangle \tag{3.9}$$

and the energy is given by

$$E = \frac{\langle\Theta|\hat{H}|\Theta\rangle}{\langle\Theta|\Theta\rangle} \tag{3.10}$$

Now if $|\Theta\rangle$ is normalized, and we substitue the expansion over exact eigenfunctions (equation 3.9) into the equation above, we obtain

$$E = \sum_i c_i^* c_i \mathcal{E}_i \tag{3.11}$$

where $\mathcal{E}_i$ is the ith energy eigenvalue, i.e. $\hat{H}|\Psi_i\rangle = \mathcal{E}_i|\Psi_i\rangle$. Now subtract $\mathcal{E}_0$, the exact nonrelativistic ground state energy, from both sides to obtain

$$E - \mathcal{E}_0 = \sum_i c_i^* c_i \mathcal{E}_i - \mathcal{E}_0 \tag{3.12}$$

or

$$E - \mathcal{E}_0 = \sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0) \tag{3.13}$$

since $|\Theta\rangle$ is normalized and $\sum_i c_i^* c_i = 1$. Since $\mathcal{E}_i$ are greater than or equal to $\mathcal{E}_0$ for all values of i and the coefficients $c_i^* c_i$ are necessarily non-negative, the right hand side of equation (3.12) is non-negative, so that $E \geq \mathcal{E}_0$. We should also point out that the variational theorem holds for the Hartree-Fock method as well as for CI, since equation (3.10) is valid for the Hartree-Fock energy—for

a given set of MO's, the HF energy can be formulated as a (trivial) 1 x 1 CI eigenvalue problem. In a similar manner, the MCSCF method (where MO's *and* CI coefficients are optimized) is also "variational."

It should be clear that instead of using the exact nonrelativistic eigenfunctions $|\Psi_i\rangle$ in equation (3.9), we could also have used an expansion over the exact eigenfunctions *within* the one-electron space spanned by $|\Theta\rangle$ (i.e. we could expand the approximate CI wavefunction $|\Theta\rangle$ in terms of the full CI wavefunctions). This means that not only is the approximate energy an upper bound to the exact nonrelativistic ground-state energy, but it is also an upper bound to the full CI energy in the given one-electron basis.

3.3   Why are Coupled-Cluster and MBPT Energies not Variational?

Electron correlation methods other than CI may not be variational. For example, consider the coupled-cluster energy expression

$$E = \frac{\langle \Phi_0 | e^{-\hat{T}} H e^{\hat{T}} | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} \qquad (3.14)$$

If the operator $e^{\hat{T}}$ is not trunctated, then we know that $e^{\hat{T}} | \Phi_0 \rangle = |\Psi_0\rangle$. Generally, however, the operator is truncated. Let us define $e^{\hat{T}'} | \Phi_0 \rangle = |\Theta_A\rangle$ for our truncated $\hat{T}'$. Now define $\langle \Phi_0 | e^{-\hat{T}'} = \langle \Theta_B |$. Note that in general $|\Theta_B\rangle \neq |\Theta_A\rangle$, which *would* have occured had we used $\langle \Phi_0 | (e^{\hat{T}'})^{\dagger}$ on the left. Then the energy expression is

$$E = \frac{\langle \Theta_B | H | \Theta_A \rangle}{\langle \Phi_0 | \Phi_0 \rangle} \qquad (3.15)$$

which, after expansion over the complete set of eigenvectors, becomes

$$E = \frac{\sum_{ij} c_i^* d_j \langle \Psi_i | H | \Psi_j \rangle}{\langle \Phi_0 | \Phi_0 \rangle} \qquad (3.16)$$

This simplifies to

$$E = \frac{\sum_i c_i^* d_i \mathcal{E}_i}{\langle \Phi_0 | \Phi_0 \rangle} \qquad (3.17)$$

At this point we can go no farther, because the terms $c_i^* d_i$ may be negative, in contrast to the situation in equation (3.12).

For completeness, we also show that MBPT energies are not variational. The nth order MBPT wavefunction may be written [12] as

$$|\Theta_{\mathrm{MBPT}}^{(n)}\rangle = |\Phi_0\rangle + \sum_{k=1}^{n} \left[\frac{\mathcal{V}(1 - |\Phi_0\rangle\langle\Phi_0|)}{E_0 - H_0} \right]^k |\Phi_0\rangle_L \tag{3.18}$$

where the sum is over "linked diagrams" only. The nth order energy is then given by

$$E_{\mathrm{MBPT}}^{(n)} = \langle\Phi_0|\hat{H}|\Theta_{\mathrm{MBPT}}^{(n-1)}\rangle \tag{3.19}$$

Since this integral is not symmetric, the energy is not variational. Only the first-order perturbation theory energy (which is also the Hartree-Fock energy) is variational, since it uses $|\Theta_{\mathrm{MBPT}}^{(0)}\rangle = |\Phi_0\rangle$.

3.4   Application of the Variational Theorem to Other States

In this section, we parallel the arguments of Pauling and Wilson [13], p. 186. So far, we have shown only that the energy calculated as the expectation value of some trial function must be an upper bound to the true ground state energy $\mathcal{E}_0$. In certain cases, we may derive a similar result for other states. If we take a trial function $|\Phi_0\rangle$ such that the first k coefficients in equation (3.9) are zero, then we may subtract $\mathcal{E}_k$ from equation (3.11) to obtain

$$E - \mathcal{E}_k = \sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_k) \tag{3.20}$$

Now since we've assumed $c_j = 0, j = 0, 1, \ldots, k$, this simplifies to

$$E - \mathcal{E}_k = \sum_{i=k} c_i^* c_i (\mathcal{E}_i - \mathcal{E}_k) \tag{3.21}$$

Once again, we can see that every term on the right side is nonnegative, so $E - \mathcal{E}_k \geq 0$.

There are any number of cases in which we have a trial function of the form just described. Consider, for example, a calculation on a triplet state for a molecule which has a singlet ground state. If our trial function is constrained to be a triplet, then all singlet eigenfunctions $|\Psi_i\rangle$ will have zero coefficients in the expansion of the trial function. In this case, the energy we minimize from the triplet trial function will be an upper bound to the lowest triplet energy, even though there is a lower-lying singlet state. Similar arguments can be made for spatial symmetry.

3.5 Convergence of the Wavefunction

A consequence of the variational theorem is that as the energy E of an approximate variational wavefunction approaches the exact energy $\mathcal{E}_0$, the approximate wavefunction $|\Phi\rangle$ approaches the exact one $|\Psi_0\rangle$. This follows from equation (3.12), which shows that as the energy E is minimized (or, equivalently, as $E - \mathcal{E}_0$ is minimized), then $\sum_i c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0)$ is minimized; that is, the sum of squares of the absolute values of the coefficients of excited states with weight factors $(\mathcal{E}_i - \mathcal{E}_0)$ is minimized. It is apparent that these weight factors might not be optimal if we want the $|\Phi\rangle$ which gives the best value for a property other than the energy, such as dipole moment. However, in the limit that E is minimized with a sufficiently large basis so that $E = \mathcal{E}_0$, then $\sum_{i=1} c_i^* c_i (\mathcal{E}_i - \mathcal{E}_0) = 0$, or $c_0 = 1$, implying that $|\Phi\rangle = |\Psi_0\rangle$. Once we have the exact wavefunction, then of course all properties can be computed exactly (again, within the Born-Oppernheimer approximation and neglecting relativistic effects). To summarize, variational improvements in the energy give improvements in the approximate wavefunction, which in turn improves the values of all other properties; however, these other properties will not necessarily converge as fast as the energy with respect to N-electron basis set improvement.

3.6   Variational Theorem Bounds on Excited States

Just as the lowest eigenvalue has been shown to be an upper bound to the exact ground-state energy, more generally, any eigenvalue E_i can be shown to be an upper bound to the corresponding exact excited state energy $\mathcal{E}_i$ [14]. In fact, one can also show that as other N-electron basis functions are added to the CI procedure, the eigenvalues obey the MacDonald-Hylleraas-Undheim relations [14, 15]

$$E_{i-1}^{(m)} \leq E_i^{(m+1)} \leq E_i^{(m)} \tag{3.22}$$

where m is the number of N-electron basis functions.

4   Reducing the Size of the CI Space

4.1   Symmetry Restrictions on the CI Space

In this section we discuss some general considerations concerning which N-electron basis functions should be included in the CI space (given that, in the general case, there are too many for us to include all of them). Certainly if we can find a class of N-electron functions which rigorously have zero Hamiltonian matrix elements with the desired CI wavefunction, then none of these basis functions will contribute at all to our approximate wavefunction and they should not be included in the CI space; the Hamiltonian will be block diagonal, and all the functions in this class will contribute to the wrong block. We will now prove the following: consider an operator $\hat{A}$ which commutes with $\hat{H}$. If

$$\hat{A}|\Phi_1\rangle = a_1|\Phi_1\rangle \tag{4.1}$$

and

$$\hat{A}|\Phi_2\rangle = a_2|\Phi_2\rangle, a_1 \neq a_2 \tag{4.2}$$

then

$$\langle\Phi_1|\hat{H}|\Phi_2\rangle = 0 \tag{4.3}$$

First we show that $\hat{H}|\Phi_2\rangle$ is an eigenfunction of $\hat{A}$ with eigenvalue a_2. Define

$$\hat{H}|\Phi_2\rangle = |\Phi_2'\rangle \tag{4.4}$$

Now apply $\hat{A}$ to $|\Phi_2'\rangle$

$$
\begin{aligned}
\hat{A}\left[\hat{H}|\Phi_2\rangle\right] &= \hat{A}\hat{H}|\Phi_2\rangle \\
&= \hat{H}\hat{A}|\Phi_2\rangle \\
&= \hat{H}a_2|\Phi_2\rangle \\
&= a_2\left[\hat{H}|\Phi_2\rangle\right]
\end{aligned}
\tag{4.5}
$$

Where we have used the given that $\hat{A}\hat{H} = \hat{H}\hat{A}$. We may now write

$$\hat{A}|\Phi_2'\rangle = a_2|\Phi_2'\rangle \tag{4.6}$$

Now consider again equation (4.1). If we take the adjoint of this equation we obtain

$$\langle \Phi_1 | \hat{A}^\dagger = \langle \Phi_1 | a_1^* \tag{4.7}$$

Now use the fact that $\hat{A} = \hat{A}^\dagger$ (we assumed $\hat{A}$ was Hermitian) and that the eigenvalues of a Hermitian operator are real. This yields

$$\langle \Phi_1 | \hat{A} = \langle \Phi_1 | a_1 \tag{4.8}$$

Multiply on the right by $|\Phi_2'\rangle$

$$\langle \Phi_1 | \hat{A} | \Phi_2' \rangle = a_1 \langle \Phi_1 | \Phi_2' \rangle \tag{4.9}$$

Now multiply equation (4.6) on the left by $\langle \Phi_1 |$ to obtain

$$\langle \Phi_1 | \hat{A} | \Phi_2' \rangle = a_2 \langle \Phi_1 | \Phi_2' \rangle \tag{4.10}$$

If we subtract equation (4.10) from equation (4.9) we arrive at

$$(a_1 - a_2)\langle \Phi_1 | \Phi_2' \rangle = 0 \tag{4.11}$$

Since we assumed $a_1 \neq a_2$, then $\langle \Phi_1 | \Phi_2' \rangle = 0$. Recalling the definition of $|\Phi_2'\rangle$, we have

$$\langle \Phi_1 | \hat{H} | \Phi_2 \rangle = 0 \tag{4.12}$$

which was to be proven. Thus if our desired wavefunction is an eigenfunction of some Hermitian operator that commutes with the Hamiltonian, our CI space need not include those N-electron functions which are eigenfunctions of this operator with different eigenvalues. As an example, consider the spin angular momentum operator $\hat{S}^2$. If we want to solve for a state $|\Psi\rangle$ of spin S, then we know

$$\hat{S}^2 |\Psi\rangle = S(S+1)|\Psi\rangle \tag{4.13}$$

and any basis function of a different spin can be excluded from the CI. Slater determinants are generally not eigenfunctions of $\hat{S}^2$. However, we can take linear combinations of Slater determinants so that we do have eigenfunctions of $\hat{S}^2$; such functions are generally called configuration state functions, or CSF's. The advantage of using CSF's is that we can throw out all functions with the wrong

eigenvalue S—they contribute to another, noninteracting block of the Hamiltonian matrix. This reasoning also applies to symmetry operations of point groups, such that we can throw away any N-electron basis functions (whether determinants or CSF's) which have the wrong irreducible representation. We can also restrict the basis functions according to their eigenvalues with respect to the operator S_z. For a triplet state, we can perform the calculation using basis functions which have $M_s = -1, 0$, or 1. If we were to include basis functions of *all* these values of M_s, we would obtain a triply-degenerate answer—as one should expect![3]

4.2   Classification of Basis Functions by Excitation Level

Now we will discuss the importance of various excitation classes to the CI wavefunction. As noted in equation (2.9), the CI expansion is typically truncated according to excitation level; in the vast majority of CI studies, the expansion is truncated (for computational tractability) at doubly-excited configurations. Since the Hamiltonian contains only two-body terms, only singles and doubles can interact directly with the reference (for the sake of simplicity, we are assuming only a single reference for now). This is a direct result of Slater's Rules (cf. section 2.4). The structure of the CI matrix with respect to excitation level is given below (adapted from Szabo and Ostlund [1], p. 235), where $|S\rangle, |D\rangle, |T\rangle$, and $|Q\rangle$ represent blocks of singly, doubly, triply, and quadruply excited determinants, respectively. The Hamiltonian matrix $\mathbf{H}$ is Hermitian; if only real orbitals are used, as is usually the case, then the Hamiltonian is also symmetric. Thus only the lower triangle of $\mathbf{H}$ is shown below.

$$
\mathbf{H} = \begin{array}{c} \langle\Phi_0| \\ \langle S| \\ \langle D| \\ \langle T| \\ \langle Q| \\ \vdots \end{array}
\left[\begin{array}{ccccccc}
\langle\Phi_0|H|\Phi_0\rangle & & & & & & \cdots \\
0 & \langle S|H|S\rangle & & & & & \cdots \\
\langle D|H|\Phi_0\rangle & \langle D|H|S\rangle & \langle D|H|D\rangle & & & & \cdots \\
0 & \langle T|H|S\rangle & \langle T|H|D\rangle & \langle T|H|T\rangle & & & \cdots \\
0 & 0 & \langle Q|H|D\rangle & \langle Q|H|T\rangle & \langle Q|H|Q\rangle & & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array} \right] \quad (4.14)
$$

---

[3]However, state-of-the-art determinant-based CI algorithms often include computational simplifications if the $M_s = 0$ component is used [16].

Note that the matrix elements $\langle S|H|\Phi_0\rangle$ are given as 0. This is due to Brillouin's theorem, which is valid when our reference function $|\Phi_0\rangle$ is obtained by the Hartree-Fock method (Hartree-Fock guarantees that off-diagonal elements of the Fock matrix are zero, and it turns out that the matrix element between two Slater determinants which differ by one spin orbital is equal to an off-diagonal element of the Fock matrix). Furthermore, the blocks $\langle X|H|Y\rangle$ which are not necessarily zero may still be sparse; for example, the matrix element $\langle \Phi_{ab}^{rs}|H|\Phi_{cdef}^{tuvw}\rangle$, which belongs to the block $\langle D|H|Q\rangle$, will be nonzero only if a and b are contained in the set $\{c,d,e,f\}$ and if r and s are contained in the set $\{t,u,v,w\}$.

Since only the doubles interact directly with the Hartree-Fock reference, we expect double excitations to make the largest contributions to the CI wavefunction, after the reference state. Indeed, this is what is observed. Even though singles, triples, etc. do not interact directly with the reference, they can still become part of the CI wavefunction (i.e. have non-zero coefficients) because they mix with the doubles, directly or indirectly. Although singles are much less important to the energy than doubles, they are generally included in CI treatments because of their relatively small number and because of their greater importance in describing one-electron properties (dipole moment, etc.)

4.3 Energy Contributions of the Various Excitation Levels

Table 2 demonstrates the importance of various excitation classes in obtaining CI energies. We see that singles and doubles account for 95% of the correlation energy at the equilibrium geometries of the molecules listed. We see that quadruple excitations are more important than triples, at least as far as the energy is concerned. At stretched geometries, the CISD and CISDT methods become markedly poorer, yet the CISDTQ method still recovers a very high (and nearly constant) fraction of the correlation energy, suggesting that CISDTQ should give reliable results for energy differences across potential energy surfaces for molecules of this size.

Table 2: Percentage of correlation energy recovered by various CI excitation levels for some small molecules.

Molecule	Percent Corr. Energy[a]		
	CISD	CISDT	CISDTQ
BH	94.91	n/a	99.97
$H_2O(R_e)$	94.70	95.47	99.82
$H_2O(1.5\ R_e)$	89.39	91.15	99.48
$H_2O(2.0\ R_e)$	80.51	83.96	98.60
NH_3	94.44	95.43	99.84
HF	95.41	96.49	99.86
H_7^+	96.36	96.87	99.96

[a]Data from reference [6] except for H_7^+ data from reference [17].

4.4   Size of the CI Space as a Function of Excitation Level

We can also see from Table 3 that the number of N-electron basis functions increases dramatically with increasing excitation level. It should be pointed out that while the calculations on BH, HF, and H_7^+ used DZP basis sets, those on H_2O and NH_3 used only DZ basis sets. A DZP basis should be considred the minimum adequate basis for a truly meaningful benchmark study, and even then it is desirable to use a high-quality basis such as an Atomic Natural Orbital (ANO) set. While it is generally possible to perform CISD calculations on small molecules with a good one-electron basis, the CISDTQ method is limited to molecules containing very few heavy atoms, due to the extreme number of N-electron basis functions required. Full CI calculations are of course even more difficult to perform, so that despite their importance as benchmarks, few full CI energies using flexible one-electron basis sets have been obtained.

The size of the full CI space in CSF's can be calculated (including spin symmetry but ignoring spatial symmetry) by Weyl's dimension formula as applied to the Distinct row table (DRT). If N is the number of electrons, n is the number of orbitals, and S is the total spin, then the dimension of the CI space in CSF's

Table 3: Number of CSF's required for small molecules at several levels of CI.

Molecule	CISD	CISDT	CISDTQ	FCI
	CSF's required[a]			
BH	568	n/a	28 698	132 686
H_2O	361	3 203	17 678	256 473
NH_3	461	4 029	19 925	137 321
HF	552	6 712	48 963	944 348
H_7^+	1 271	24 468	248 149	2 923 933

[a]Data from reference [6] except for H_7^+ data from reference [17].

is given by

$$D_{nNS} = \frac{2S+1}{n+1} \begin{pmatrix} n+1 \\ N/2 - S \end{pmatrix} \begin{pmatrix} n+1 \\ N/2 + S + 1 \end{pmatrix} \tag{4.15}$$

The dimension of the full CI space in determinants (again, ignoring spatial symmetry) is computed simply by

$$D_{nN_\alpha N_\beta} = \begin{pmatrix} n \\ N_\alpha \end{pmatrix} \begin{pmatrix} n \\ N_\beta \end{pmatrix} \tag{4.16}$$

or, in a form closer to equation 4.15,

$$D_{nNS} = \begin{pmatrix} n \\ N/2 + S \end{pmatrix} \begin{pmatrix} n \\ N/2 - S \end{pmatrix} \tag{4.17}$$

Table 4 shows the dimension of the full CI space (neglecting spatial symmetry) in determinants and in CSF's. Current full CI algorithms are limited to a few million determinants. Although there have been reports [18, 19, 20] of larger calculations (including a few billion determinants), the computational expense is (currently) too great for routine calculations of this size.

Table 4: Dimension of Full CI in Determinants (CSF's in parentheses)

Orbitals	Number of electrons			
	6	8	10	12
10	14.4×10^3	44.1×10^3	63.5×10^3	44.1×10^3
	(4.95×10^3)	(13.9×10^3)	(19.4×10^3)	(13.9×10^3)
20	1.30×10^6	23.5×10^6	240×10^6	1.50×10^9
	(379×10^3)	(5.80×10^6)	(52.6×10^6)	(300×10^6)
30	16.5×10^6	751×10^6	20.3×10^9	353×10^9
	(4.56×10^6)	(172×10^6)	(4.04×10^9)	(62.5×10^9)

4.5   The Frozen Core Approximation

It is quite common in applications of the CI method to invoke the frozen core approximation, in which the lowest-lying molecular orbitals (occupied by the inner-shell electrons) are constrained to remain doubly-occupied in all configurations. The frozen core for atoms lithium to neon typically consists of the 1s atomic orbital, while that for atoms sodium to argon consists of the atomic orbitals 1s, 2s, $2p_x$, $2p_y$ and $2p_z$. The frozen molecular orbitals are those which are primarily these inner-shell atomic orbitals (or linear combinations thereof).

A justification for this approximation is that the inner-shell electrons of an atom are less sensitive to their environment than are the valence electrons. Thus the error introduced by freezing the core orbitals is nearly constant for molecules containing the same types of atoms. In fact, it is sometimes *recommended* that one employ the frozen core approximation as a general rule because most of the basis sets commonly used in quantum chemical calculations do not provide sufficient flexibility in the core region to accurately describe the correlation of the core electrons.

Not only does the frozen core approximation reduce the number of configurations in the CI procedure, but it also reduces the computational effort required

to evaluate matrix elements between the configurations which remain. Assuming that all frozen core orbitals are doubly occupied and orthogonal to all other molecular orbitals, then it can be shown [21] that

$$\langle \Phi_I | \hat{H} | \Phi_J \rangle = \langle \bar{\Phi}_I | \hat{H}_0 | \bar{\Phi}_J \rangle \tag{4.18}$$

where $\bar{\Phi}_I$ and $\bar{\Phi}_J$ are identical to Φ_I and Φ_J, respectively, except that the core orbitals have been deleted from $\bar{\Phi}_I$ and $\bar{\Phi}_J$, and $\hat{H}$ has been replaced by $\hat{H}_0$ defined by

$$\hat{H}_0 = E_c + \sum_{i=1}^{N-N_c} \hat{h}_c(i) + \sum_{i>j}^{N-N_c} \frac{1}{r_{ij}} \tag{4.19}$$

where N is the number of electrons and N_c is the number of core electrons. E_c is the so-called "frozen-core energy," which is the expectation value of the determinant formed from only the N_c core electrons doubly occupying the $n_c = N_c/2$ core orbitals

$$E_c = 2 \sum_i^{n_c} h_{ii} + \sum_{ij}^{n_c} \{2(ii|jj) - (ij|ji)\} \tag{4.20}$$

Finally, $\hat{h}_c(i)$ is the one-electron Hamiltonian operator for electron i in the average field produced by the N_c core electrons,

$$\hat{h}_c(i) = \hat{h}(i) + \sum_{j=1}^{n_c} \left\{ 2\hat{J}_j(i) - \hat{K}_j(i) \right\} \tag{4.21}$$

with $\hat{J}_j(i)$ and $\hat{K}_j(i)$ representing the standard Coulomb and exchange operators, respectively. Note that, although we have written the frozen core energy E_c and frozen core operator $\hat{h}_c$ in terms of molecular orbitals, it is not necessary to explicitly transform the one- and two-electron integrals involving core orbitals. Assuming real orbitals, we can define a frozen core density matrix [22] in atomic (or symmetry) orbitals as

$$P_{\rho\sigma}^c = \sum_i^{n_c} C_\rho^i C_\sigma^i \tag{4.22}$$

where C_ρ^i is the contribution of atomic orbital ρ to molecular orbital i. Now the frozen core operator in atomic orbitals becomes

$$h_{\mu\nu}^c = h_{\mu\nu} + 2 \sum_{\rho\sigma} (\rho\sigma|\mu\nu) P_{\rho\sigma}^c - \sum_{\rho\sigma} (\rho\mu|\nu\sigma) P_{\rho\sigma}^c \tag{4.23}$$

and the frozen core operator in molecular orbitals h_{ij}^c can be obtained simply by transforming $h_{\mu\nu}^c$. Similarly the frozen core energy can be evaluated as

$$
\begin{aligned}
E_c &= \sum_{\mu\nu} P_{\mu\nu}^c \left(h_{\mu\nu} + h_{\mu\nu}^c \right) \\
&= Tr(P^c h) + Tr(P^c h^c)
\end{aligned}
\tag{4.24}
$$

An analogous approximation is the *deleted virtual* approximation, whereby a few of the highest-lying virtual (unoccupied) molecular orbitals are constrained to remain unoccupied in all configurations. Since these orbitals can never be occupied, they can be removed from the CI procedure entirely because no terms involving them contribute to the CI coefficients or energy. The rationalization for this procedure is that it is unlikely that electrons will choose to partially populate high-energy orbitals in their attempt to avoid other electrons. However, this conclusion is generally true only for very high-lying virtual orbitals (such as those formed by antisymmetric combinations of symmetry orbitals in the core region). For all other virtual orbitals, such simplistic reasoning is not sufficient.

Davidson points out that those high energy SCF virtual orbitals which result from the antisymmetric combination of the two basis functions describing each atomic orbital in a double-ζ basis set (such as the 3p-like orbital formed from the minus combination of the larger and smaller 2p atomic orbitals on oxygen) often make the largest contribution to the correlation energy in Møller-Plesset (MPn) wavefunctions [23]. This can be seen from the expression for the second-order correction to the energy in Møller-Plesset perturbation theory,

$$
E^{(2)} = \sum_{a>b} \sum_{r>s} \frac{|\langle ab||rs \rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}
\tag{4.25}
$$

where ϵ_i is the orbital energy (eigenvalue) of orbital i. Thus a virtual orbital r with a large energy ϵ_r will contribute to a large energy denominator in each term of equation (4.25) in which it appears. However, if orbital r lies close spatially to one of the orbitals a or b occupied in the reference, then this large overlap will contribute to a large two-electron integral $\langle ab||rs \rangle$. This integral is *squared* in

the numerator, leading to a large overall contribution to the second-order energy. For the antisymmetric combinations of the basis functions describing the *core region*, this large numerator is insufficient to overcome the even larger energy denominator; such virtual orbitals can generally be deleted with minimal loss in the correlation energy recovered.

Although the analysis in the preceeding paragraph is based on perturbation theory, similar conclusions can be drawn for the CI method. This is most easily verified by actual calculations, since analytical expressions for the energetic contribution of orbitals to the CI energy are not nearly as simple to obtain or interpret as is equation (4.25).

4.6 Truncated CI is not Size Extensive

As we have previously pointed out, full CI—being the matrix formulation of the Schrödinger equation—is an *exact* theory for nonrelativistic electronic structure problems. If we truncate the CI (either in the one-electron or N-electron space), we no longer have an exact theory. Of course either of these truncations will introduce an error in the wavefunction, which will cause errors in the energy and all other properties. One particularly unwelcome result of truncating the N-electron basis is that the CI energies obtained are no longer size extensive or size consistent.

These two terms, size extensive and size consistent, are used somewhat loosely in the literature. Of the two, size extensivity is the most well-defined. A method is said to be size extensive if the energy calculated thereby scales linearly with the number of particles N. The word "extensive" is used in the same sense as in thermodynamics, when we refer to an *extensive*, rather than an *intensive* property. A method is called size consistent if it gives an energy $E_A + E_B$ for two well separated subsystems A and B. While the definition of size extensivity applies at any geometry, the concept of size consistency applies only in the limiting case of infinite separation. In addition, size consistency *usually* also implies correct dissociation into fragments; this is the source of much of the confusion arising from this term. Thus restricted Hartree-Fock (RHF) is size extensive, but it is not necessarily size consistent, since it cannot properly describe dissociation into open-shell fragments. It can be shown that many-body perturbation theory (MBPT) and coupled-cluster (CC) methods are size extensive, but they will be size consistent only if they are based on reference wavefunction which dissociates properly.

As previously stated, truncated CI's are neither size extensive nor size consistent. A simple (and often used!) example is sufficient to make the point. Consider two noninteracting hydrogen molecules. If the CISD method is used, then the energy of the two molecules at large separation will not be the same as the sum of their energies when calculated separately. In order for this to be the case,

we would have to include *quadruple* excitations in the supermolecule calculation, since local double excitations could happen simultaneously on A and B.

We would tend to think that size extensivity and size consistency are important, physical properties that all quantum mechanical models should have (indeed, full CI, an exact theory, has these properties), but perhaps they are not as essential as all that. Duch and Diercksen have claimed that "making size extensivity the most important requirement of quantum chemical methods, although it does not guarantee correct physical description, seems to be based not that much on physical as on esthetical criteria" [24]. Indeed, they show that quantum mechanics is a "holistic" theory, not well-suited toward the description of separated subsystems:

> Hilbert space of antisymmetric, many particle functions, describing the total system, can not be decomposed into separate subspaces. Consider two systems, S_A and S_B, with N_A and N_B electrons, respectively. Each system is described by its own function, Ψ_A antisymmetric in N_A particles and Ψ_B in N_B. Assuming that both functions are normalized to unity it is easy to show that the product function $\Psi_{AB} = \Psi_A \Psi_B$ is always "far" from the antisymmetric function $\Psi = \mathcal{A}\Psi_{AB}$, as measured by the overlap $\langle \Psi_{AB} | \Psi \rangle$ or the norm of the difference $2 - \sqrt{2} \leq ||\Psi_{AB} - \Psi||^2 \leq 2$.

Such arguments notwithstanding, it is clear that the fraction of the correlation energy recovered by a truncated CI will diminish as the size of the system increases, making it a progressively less accurate method. There have been many attempts to correct the CI energy to make it size extensive. The most widely-used (and simplest) of these methods is referred to as the Davidson correction [25], which is

$$\Delta E_{DC} = E_{SD}(1 - c_0^2) \tag{4.26}$$

where E_{SD} is the basis set correlation energy recovered by a CISD procedure. This correction approximately accounts for the effects of "unlinked quadruple" excitations (i.e. simultaneous pairs of double excitations). The multireference version [24] of this correction is

$$\Delta E_{DC} = \left(1 - \sum_{i \in \text{Ref}} |c_i|^2\right)(E_{MRCI} - E_{MR}) \tag{4.27}$$

where E_{MRCI} is the multireference CI energy and E_{MR} is the energy obtained from the set of references (MCSCF energy if the references are obtained as all references in an MCSCF procedure). We have simply replaced the CISD correlation energy in equation (4.26) with the analogous multireference correlation energy, and we have replaced c_0^2 with the analogous sum of squares of all the reference coefficients.

There are a number of other size extensivity corrections, and most of them do not take any significant amount of computation. Reference [24] provides a nice comparison of several of the more common CI size extensivity correction methods. We should also mention that Malrieu and co-workers have presented a self-consistent dressing of the Hamiltonian which gives size extensive results for selected CI procedures [26].

5   Second Quantization

Much of the literature in CI theory makes use of the notation of second-quantization. Szabo and Ostlund [1] give a good introduction to second-quantized operators. Here we will only summarize the anticommutation relations between creation and annihilation operators, and then proceed to express the Hamiltonian in second quantized form for *spatial* orbitals, rather than for spin orbitals. Then we will use these results to derive the Hamiltonian in terms of the unitary group generators.

The anticommutation relations for two annihilation operators is

$$\{a_j, a_i\} = a_j a_i + a_i a_j = 0 \tag{5.1}$$

and the anticommutation relation for two creation operators is similarly

$$\{a_j^\dagger, a_i^\dagger\} = a_j^\dagger a_i^\dagger + a_i^\dagger a_j^\dagger = 0 \tag{5.2}$$

The anticommutation relation between a creation and an annihilation operator is

$$\{a_i, a_j^\dagger\} = a_i a_j^\dagger + a_j^\dagger a_i = \delta_{ij} \tag{5.3}$$

Now we will find an expression for the Hamiltonian in terms of creation and annihilation operators over spatial orbitals. We begin with the second-quantized form of the one- and two-electron operators (see Szabo and Ostlund [1] p. 95)

$$\hat{\mathcal{O}}_1 = \sum_{ij}^{2n} \langle i|h|j\rangle a_i^\dagger a_j \tag{5.4}$$

$$\hat{\mathcal{O}}_2 = \frac{1}{2} \sum_{ijkl}^{2n} \langle ij|kl\rangle a_i^\dagger a_j^\dagger a_l a_k \tag{5.5}$$

where the sums run over all spin orbitals $\{\chi_i\}$. Thus the Hamiltonian is

$$\hat{H} = \sum_{pq}^{2n} a_p^\dagger a_q [p|h|q] + \frac{1}{2} \sum_{pqrs}^{2n} a_p^\dagger a_r^\dagger a_s a_q [pq|rs] \tag{5.6}$$

From the previous equation we can see that the second-quantized form of the Hamiltonian is independent of the number of electrons in the system. Now integrate over spin, assuming that spatial orbitals are constrained to be identical for

α and β spins. A sum over all $2n$ spin orbitals can be split up into two sums, one over n orbitals with α spin, and one over n orbitals with β spin. Symbolically, this is

$$\sum_a^{2n} = \sum_a^n + \sum_{\bar{a}}^n \tag{5.7}$$

The one-electron part of the Hamiltonian becomes

$$\hat{H}_{\text{one}} = \sum_{pq}^n [p|h|q]a_{p\alpha}^\dagger a_{q\alpha} + [p|h|\bar{q}]a_{p\alpha}^\dagger a_{q\beta} + [\bar{p}|h|q]a_{p\beta}^\dagger a_{q\alpha} + [\bar{p}|h|\bar{q}]a_{p\beta}^\dagger a_{q\beta} \tag{5.8}$$

After integrating over spin, this becomes

$$\hat{H}_{\text{one}} = \sum_{pq}^n (p|h|q)\{a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}\} \tag{5.9}$$

The two-electron term can be expanded similarly to give

$$\hat{H}_{\text{two}} = \frac{1}{2}\sum_{pqrs}^n (pq|rs)\{a_{p\alpha}^\dagger a_{r\alpha}^\dagger a_{s\alpha}a_{q\alpha} + a_{p\alpha}^\dagger a_{r\beta}^\dagger a_{s\beta}a_{q\alpha} + a_{p\beta}^\dagger a_{r\alpha}^\dagger a_{s\alpha}a_{q\beta} + a_{p\beta}^\dagger a_{r\beta}^\dagger a_{s\beta}a_{q\beta}\} \tag{5.10}$$

Now we make use of the anticommutation relation (5.1) and we swap the order of a_s and a_q, introducing a minus sign. This yields

$$\hat{H}_{\text{two}} = -\frac{1}{2}\sum_{pqrs}^n (pq|rs)\{a_{p\alpha}^\dagger a_{r\alpha}^\dagger a_{q\alpha}a_{s\alpha} + a_{p\alpha}^\dagger a_{r\beta}^\dagger a_{q\alpha}a_{s\beta} + a_{p\beta}^\dagger a_{r\alpha}^\dagger a_{q\beta}a_{s\alpha} + a_{p\beta}^\dagger a_{r\beta}^\dagger a_{q\beta}a_{s\beta}\} \tag{5.11}$$

Now we use the anticommutation relation between a creation and an annihilation operator, which is given by (5.3). This relation allows us to swap the a_q and $a_r^\dagger$ in each term, to give

$$\begin{aligned}
\hat{H}_{\text{two}} &= \frac{1}{2}\sum_{pqrs}^n (pq|rs)\Big[a_{p\alpha}^\dagger a_{q\alpha}a_{r\alpha}^\dagger a_{s\alpha} - \delta_{q\alpha,r\alpha}a_{p\alpha}^\dagger a_{s\alpha} + a_{p\alpha}^\dagger a_{q\alpha}a_{r\beta}^\dagger a_{s\beta} - \delta_{q\alpha,r\beta}a_{p\alpha}^\dagger a_{s\beta} \\
&+ a_{p\beta}^\dagger a_{q\beta}a_{r\alpha}^\dagger a_{s\alpha} - \delta_{q\beta,r\alpha}a_{p\beta}^\dagger a_{s\alpha} + a_{p\beta}^\dagger a_{q\beta}a_{r\beta}^\dagger a_{s\beta} - \delta_{q\beta,r\beta}a_{p\beta}^\dagger a_{s\beta}\Big]
\end{aligned} \tag{5.12}$$

Now we observe that $\delta_{q\alpha,r\alpha}$ and $\delta_{q\beta,r\beta}$ can both be written δ_{qr}, and also that $\delta_{q\alpha,r\beta}$ and $\delta_{q\beta,r\alpha}$ are both 0. This simplifies our equation to

$$\hat{H}_{\text{two}} = \frac{1}{2}\sum_{pqrs}^n (pq|rs)\Big[a_{p\alpha}^\dagger a_{q\alpha}a_{r\alpha}^\dagger a_{s\alpha} + a_{p\alpha}^\dagger a_{q\alpha}a_{r\beta}^\dagger a_{s\beta} + a_{p\beta}^\dagger a_{q\beta}a_{r\alpha}^\dagger a_{s\alpha} + a_{p\beta}^\dagger a_{q\beta}a_{r\beta}^\dagger a_{s\beta}$$

$$- \quad \delta_{qr} a_{p\alpha}^{\dagger} a_{s\alpha} - \delta_{qr} a_{p\beta}^{\dagger} a_{s\beta} \Big] \tag{5.13}$$

Now we introduce the replacement (or shift) operator

$$\hat{E}_{ij} = a_{i\alpha}^{\dagger} a_{j\alpha} + a_{i\beta}^{\dagger} a_{j\beta} \tag{5.14}$$

which Paldus has shown [5] to be isomorphic to the generators of the unitary group. This replacement operator is commonly referred to as a unitary group generator, but as Duch has pointed out [27], such usage is somewhat dubious in papers where no unitary group theory is employed.

$$\hat{H} = \sum_{pq}^{n} (p|h|q) \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs}^{n} (pq|rs) \left(\hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} \right) \tag{5.15}$$

This is the Hamiltonian in terms of replacement operators. Contemporary papers on CI theory often express the Hamiltonian in the form of equation (5.15).[4]

---

[4]Reference [28] contains a minor mistake, giving $(pq|rs)$ as $\langle pq|rs \rangle$.

6 Determinant-Based CI

6.1 Introduction to Determinant CI

As we have already pointed out, the size of the CI space can be reduced signif-
icantly by including only those N-electron basis functions which have the same
value of the quantum number S as the desired approximate wavefunction (cf.
sections 4.1 and 4.4). Thus it would seem that one should always prefer CSF's
to Slater determinants when performing a CI. However, certain computational
advantages arise from using determinants, as we will discuss in the next few sec-
tions. Many modern algorithms for performing large CI's (i.e. more than single
and double excitations) use determinants as the N-electron basis.

Handy's 1980 paper "Multi-Root Configuration Interaction Calculations" [29]
represented a major advance in determinant-based CI, even though the paper was
more concerned with how integrals and CI coefficients are stored than with the
computational advantages of determinants over CSF's. Handy used the Cooper-
Nesbet method for performing the CI iteration; the CI coefficients are updated
according to the formula

$$\delta c_i = \frac{r_i}{(E - H_{ii})} \tag{6.1}$$

where the $\mathbf{r}$ vector is defined as

$$r_i = \sum_j (H_{ij} - E\delta_{ij})c_j \tag{6.2}$$

Handy realized that, if determinants are used as basis functions, and particularly
if these determinants are expressed as "alpha strings" and "beta strings," then H_{ij}
(and thus $\mathbf{r}$) could be evaluated very efficiently. In order to understand Handy's
reasoning, we must first define alpha and beta strings.

6.2 Alpha and Beta Strings

Although Handy was the first to use alpha and beta strings, we will use the notation of Olsen *et al.* [16]. We define an alpha string as an ordered product of creation operators for spin orbitals with alpha spin. If I_α contains a list $\{i, j, \ldots n\}$ of the n_α occupied spin orbitals with alpha spin in determinant $|I\rangle$, then the alpha string $\alpha(I_\alpha)$ is $a_{i\alpha}^\dagger a_{j\alpha}^\dagger \ldots a_{n\alpha}^\dagger$. A beta string is defined similarly. Thus we can rewrite Slater determinant $|I\rangle$ in terms of alpha and beta strings.

$$|I\rangle = |\alpha(I_\alpha)\beta(I_\beta)\rangle = \alpha(I_\alpha)\beta(I_\beta)|\rangle \tag{6.3}$$

For instance, suppose we have the Slater determinant $|I\rangle = |\phi_{1\alpha}\phi_{2\alpha}\phi_{3\alpha}\phi_{1\beta}\phi_{2\beta}\phi_{4\beta}\rangle$. Then the alpha string $\alpha(I_\alpha)$ is given by

$$\alpha(I_\alpha) = a_{1\alpha}^\dagger a_{2\alpha}^\dagger a_{3\alpha}^\dagger \tag{6.4}$$

and the beta string is given by

$$\beta(I_\beta) = a_{1\beta}^\dagger a_{2\beta}^\dagger a_{4\beta}^\dagger \tag{6.5}$$

Note that the order of the creation operators matters; if we swap the order of two creation operators within the alpha string (or within the beta string), then we introduce a sign change (see equation 5.2). Also, acting the alpha string on the vacuum first, rather than the beta string, may introduce a minus sign, depending on the number of alpha and beta electrons. Although the order of the orbitals and whether the alpha or beta string acts first is of no real consequence, we must be sure to keep our use of alpha and beta strings consistent, or sign problems will result. In most of the literature, and in these notes, the beta string will be placed to the right of the alpha string in equations like (6.3). Further, within each string, orbitals are listed in strictly increasing order.

Handy realized the following advantages to alpha and beta strings:

1. Direct CI methods often require an index vector which points to a list of all allowed excitations from a given N-electron basis function. Using alpha and

beta strings, the index vector need not be the length of the CI vector—its size is dictated by the number of alpha or beta strings, which is approximately the square root of the number of determinants. This results from the fact that (in determinant-based CI) electrons in alpha spin-orbitals can be excited only to other alpha spin-orbitals, and electrons in beta spin-orbitals can be excited only to other beta spin-orbitals.[5]

2. To form $r(I_\alpha, I_\beta)$ in equation (6.2), all functions $|\alpha(J_\alpha)\beta(J_\beta)\rangle$ which have non-zero matrix elements with $|\alpha(I_\alpha)\beta(I_\beta)\rangle$ are generated, one at a time, with the appropriate integral being looked up and multiplied by the appropriate CI coefficient. No time is wasted considering determinants which are noninteracting, and the coefficients of the integrals are simply ± 1.

3. Efficiency is increased by realizing that all integrals which enter the expression $\langle\alpha(I_\alpha)\beta(I_\beta)|\hat{H}|\alpha(J_\alpha)\beta(I_\beta)\rangle$ (equation 6.2), where $\alpha(J_\alpha)$ differs from $\alpha(I_\alpha)$ by two orbitals, are independent of $\beta(I_\beta)$.

We will make these points more clear in our discussion of the RAS CI, which is a direct extension of Handy's observations concerning alpha and beta strings. However, at this point we will proceed to discuss the graphical representation of alpha and beta strings.

6.2.1   Graphical Representation of Alpha and Beta Strings

Any CI program requires some method for assigning a numerical code to each determinant (or configuration). Determinant-based CI's also require an addressing scheme for each alpha and beta string. Such addressing can be facilitated by using graphical representations. This section discusses the construction of simple graphs representing alpha and beta strings, and how the numerical code or address of each string can be calculated from these graphs. Our approach will be based on the work of Duch, who has described [27] the graphical representation of CI spaces in considerable detail. First, we consider the simple two-slope directed

---

[5]In the determinant CI expansion, we restrict all determinants to a single value of M_s.

graphs ("digraphs") $\mathcal{G}_2(n : n_{\alpha/\beta})$ which represent alpha or beta strings without consideration of point-group symmetry. Figure 1 presents such a digraph representing all strings with five electrons in seven orbitals (as might be appropriate for the alpha or beta strings of H_2O in a minimal basis set, neglecting point-group symmetry). Each string is represented by a "walk" on the graph, from the head (at $e = o = 0$) to the tail (at $e = n_{\alpha/\beta}$, $o = n$); the graph should always be traversed in this direction. Moving straight down from vertex (e, o) to vertex $(e, o + 1)$ indicates that orbital $o + 1$ is unoccupied in the current string, while moving down diagonally from vertex (e, o) to vertex $(e + 1, o + 1)$ indicates that orbital $o + 1$ is occupied. Each vertex on the graph is assigned a weight $x(e, o)$, and each arc connecting two vertices is assigned an arc weight $Y(e, o)$ for the arc leaving vertex (e, o). Since, in general, two different arcs can leave a given vertex, we write $Y_0(e, o)$ for the arc originating from vertex (e, o) which leaves orbital $o+1$ unoccupied, and $Y_1(e, o)$ for the arc which occupies orbital $o + 1$.[6] The index or address of a string or walk is obtained by adding weights for each arc contained in the walk, i.e.

$$I_\alpha(L^\alpha) = X(L^\alpha) + \sum_{i=0}^{n} Y_{L_i}(e_i, i) \tag{6.6}$$

where L_i is the occupation (0 or 1) of the ith arc, and (e_i, i) are the coordinates of the vertices crossed by L^α. The term $X(L^\alpha)$ gives the offset of a given graph, if more than one graph is employed. The index for a determinant is given by $I(L^\alpha, L^\beta) = I_\alpha(L^\alpha)S_\alpha + I_\beta(L^\beta)$, where S_α is the number of alpha strings. This leads us to write the CI coefficients as a rectangular matrix $C(I_\alpha, I_\beta)$. Restrictions on the CI space mean that only certain subblocks of the C matrix are allowed to be nonzero.

There are several different methods for assigning the arc weights by which we calculate the index of a string according to equation (6.6). Under the *lexical ordering* scheme, the tail (n_α, n) of an alpha string graph is assigned a weight $x = 1$. Other vertex weights are computed according to the recursive formula

$$x(e, o) = x(e + 1, o + 1) + x(e, o + 1) \tag{6.7}$$

---

[6]This differs somewhat from Duch [27], who *sometimes* uses $\bar{Y}(e, o)$ to denote the arc *entering* vertex (e, o) in reverse-lexical addressing.

Using lexical ordering, typically all arc weights $Y_0(e, o)$ are set equal to zero, and the arc weights $Y_1(e, o)$ are determined according to

$$Y_1(e, o) = x(e + 1, o + 1) + x(e + 1, o) + \cdots + x(e + 1, e + 1) \qquad (6.8)$$

Figure 1a features vertex and arc weights computed in this manner. A result of the lexical ordering scheme is that paths with a fixed upper part and an arbitrary lower part have contiguous indices. The particular choice of Y values above is appropriate if the rightmost path is to have an index of zero. The same effects can be achieved using

$$Y_1(e, o) \ = \ 0 \qquad (6.9)$$
$$Y_0(e, o) \ = \ x(e + 1, o + 1) \qquad (6.10)$$

as illustrated in Figure 1b. Any walk has the same index in Figures 1a and 1b. For instance, the walk $a_{2\alpha}^\dagger a_{3\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{7\alpha}^\dagger$ has an index of $5 + 4 + 3 + 2 + 2 = 16$ (equation 6.6) from Figure 1a, and an index of $15 + 1 = 16$ from Figure 1b.

In the so-called "reverse-lexical" ordering scheme, all upper paths for a fixed lower path have contiguous indices. Vertex weights are now determined as

$$\bar{x}(e, o) = \bar{x}(e, o - 1) + \bar{x}(e - 1, o - 1) \qquad (6.11)$$

where the overbar indicates reversed-lexical ordering. Figure 2a depicts a reversed-lexical graph with all non-occupied orbital arcs set to zero. The occupied orbital arcs are computed as

$$\bar{Y}_1(e, o) = \bar{x}(e, o) \qquad (6.12)$$

Figure 2b is the same except that now all occupied arcs have weights of zero. The non-occupied arc weights are

$$\bar{Y}_0(e, o) = \bar{x}(e, o) + \bar{x}(e + 1, o + 1) + \cdots + \bar{x}(N - 1, o + N - e - 1) \qquad (6.13)$$

Note that string indices for reverse-lexical ordering are not necessarily the same as indices for lexical ordering. For the string $a_{2\alpha}^\dagger a_{3\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{7\alpha}^\dagger$ considered previously, the index is calculated as $1 + 1 + 1 + 1 + 6 = 10$ from Figure 2a, or as $5 + 5 = 10$ from Figure 2b.

Figure 1: Alpha string graph for $n_\alpha = 5, n = 7$. Vertex weights are determined according to lexical ordering, and arc weights are given so that the rightmost path has index zero. (a) All unoccupied arc weights $Y_0(e,o)$ are zero. (b) All occupied arc weights $Y_1(e,o)$ are zero.
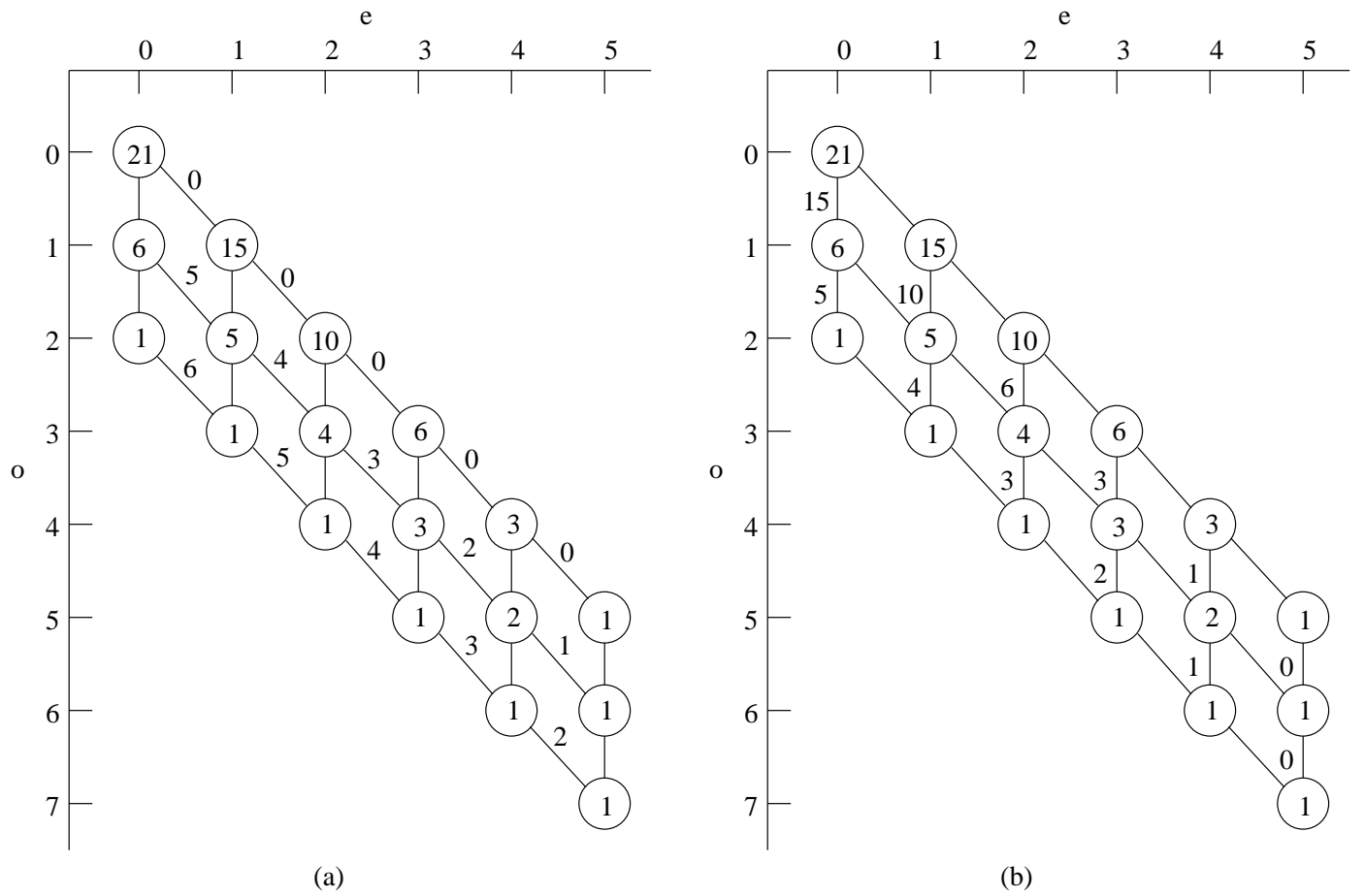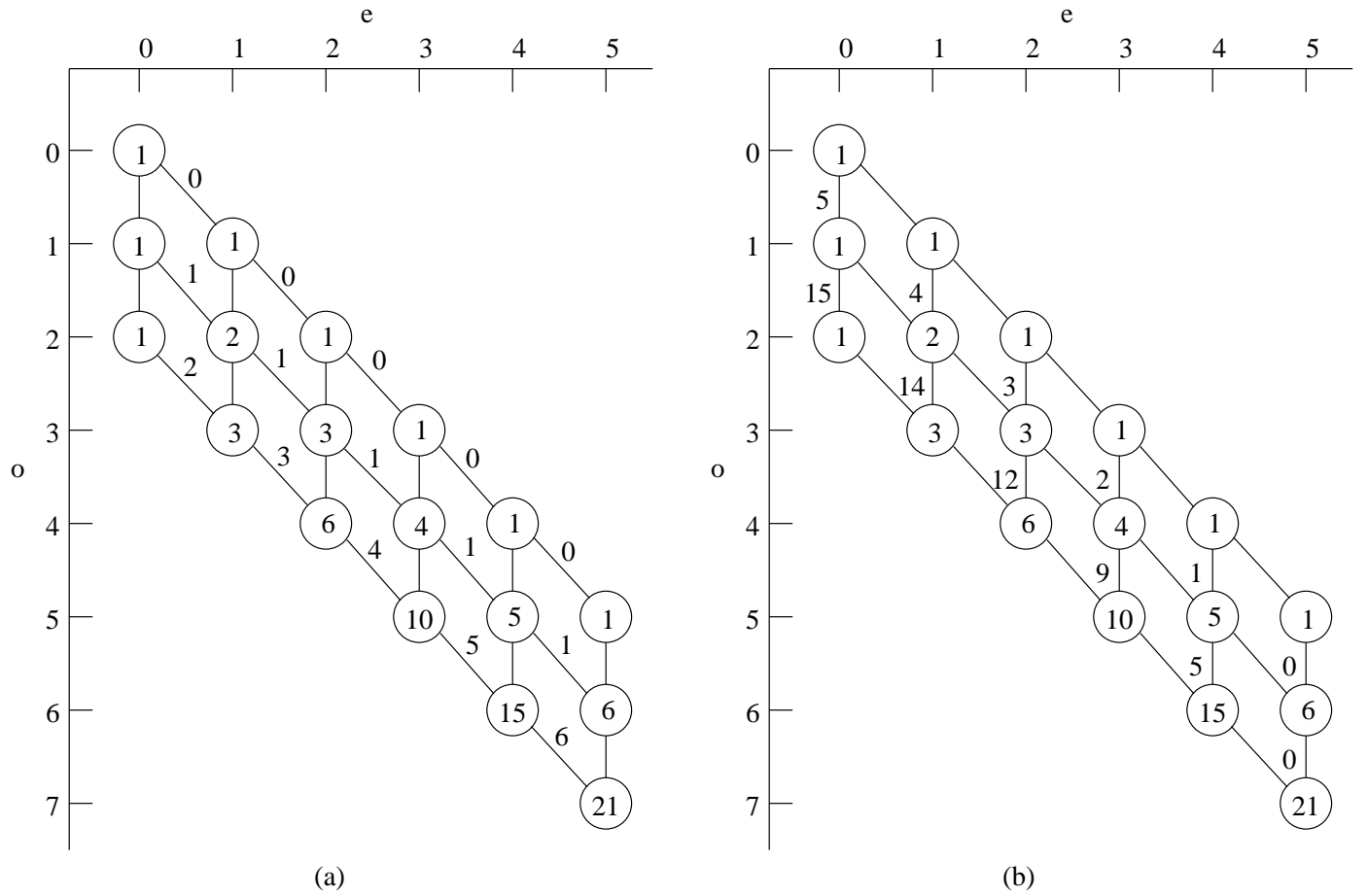

(a)   (b)

Figure 2: Alpha string graph for $n_\alpha = 5, n = 7$. Vertex weights are determined according to reverse-lexical ordering, and arc weights are given so that the rightmost path has index zero. (a) All unoccupied arc weights $Y_0(e, o)$ are zero. (b) Occupied arc weights $Y_1(e, o)$ are set to zero.

The arc weights given in Figures 1 and 2 cause the rightmost path to have an index $I(R_m) = 0$. If we change the arc weights so that the leftmost path has index $I(L_m) = 0$, we obtain four more addressing schemes. The two simplest schemes for $I(L_m) = 0$ are

$$Y_0(e, o) = 0 \quad Y_1(e, o) = x(e, o + 1) \tag{6.14}$$

$$\bar{Y}_1(e, o) = 0 \quad \bar{Y}_0(e, o) = \bar{x}(e - 1, o) \tag{6.15}$$

where the overbars indicate that reversed-lexical vertex weights have been used. Alpha strings for 5 electrons in 7 orbitals employing these addressing schemes are depicted in Figure 3.

If we add another coordinate Γ to each vertex, we can extend these simple digraphs to include point-group symmetry.

6.3   Restricted Active Space CI

The Restricted Active Space (RAS) CI method was introduced by Olsen, Roos, Jørgensen, and Aa. Jensen [16] in 1988. The RAS method calls for the partitioning of the one-electron basis into four subsets. The first subset consists of the core orbitals, which are constrained to remain doubly-occupied. The remaining three subsets are labelled I, II, and III, and the CI space is limited by requiring a *minimum of p electrons in RAS I* and a *maximum of q electrons in RAS III*. There are no restrictions on the number of electrons in RAS II, and thus it is analogous to the complete active space (CAS). The full CI can be obtained as the maximum limit of the RAS space. Interestingly, although the main focus of the paper is on the utility of the RAS method in limiting the size of CI calculations, the maximum impact of this paper has been on the development of determinant-based full CI algorithms [19, 20].

The RAS CI method relies on Handy's separation of determinants into alpha and beta strings (see section 6.2). As in other determinant-based CI methods, the basis determinants are restricted to those having a given value of M_s. Since the number of electrons N is also fixed, this means that the alpha and beta

Figure 3: Alpha string graph for $n_\alpha = 5, n = 7$, with arc weights determined so that the leftmost path has index zero. (a) Vertex weights for lexical ordering, and arc weights according to $Y_0(e,o) = 0$, $Y_1(e,o) = x(e,o+1)$. (b) Vertex weights according to reverse-lexical ordering, and arc weights according to $\bar{Y}_1(e,o) = 0$, $\bar{Y}_0(e,o) = \bar{x}(e-1,o)$.


(a)   (b)
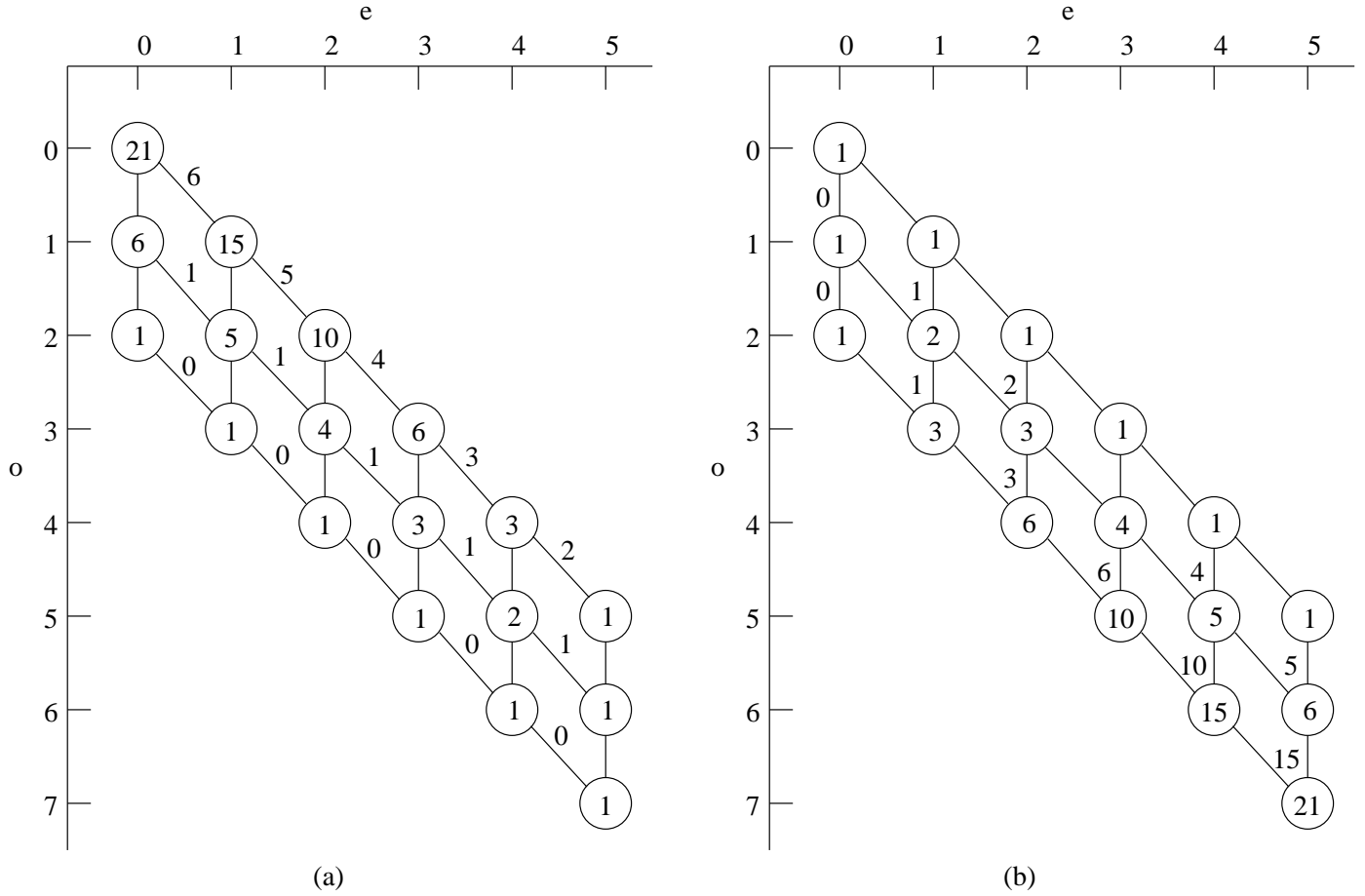
strings always have constant lengths of n_α and n_β, respectively. For a full CI, one forms all possible alpha and beta strings for a given n_α and n_β, and the basis determinants are all possible combinations of these alpha and beta strings. In a RAS CI, the CI space is restricted in two ways: first, not all alpha and beta strings are allowed, and secondly, not all combinations of alpha and beta strings to form determinants are accepted. This is best seen from an example: consider the case of 6 orbitals, with $n_\alpha = n_\beta = 3$. If orbitals 4, 5, and 6 constitute RAS III, with a maximum of 2 electrons allowed, then clearly alpha strings such as $a_{4\alpha}^\dagger a_{5\alpha}^\dagger a_{6\alpha}^\dagger$ are not allowed. Similarly, even though $a_{1\alpha}^\dagger a_{4\alpha}^\dagger a_{5\alpha}^\dagger$ and $a_{1\beta}^\dagger a_{4\beta}^\dagger a_{5\beta}^\dagger$ are allowed alpha and beta strings, these strings cannot be combined with each other because the resulting determinant would represent a quadruple excitation into RAS III.

If we employ a graphical description of the alpha and beta strings as described in section 6.2.1, in general we require one graph for alpha strings and one graph for beta strings. However, in the case that $M_s = 0$, only one graph is needed because the alpha string and beta string graphs are identical. As previously mentioned, for a RAS space not all alpha and beta strings can be freely combined. While it would be possible to create a table listing all allowed combinations of alpha and beta strings, there is a more efficient way around this difficulty. Instead of using a single graph to represent all alpha (or beta) strings, instead we use several graphs. For example, we might use one graph for all strings with no electrons in RAS III, one graph for all strings with one electron in RAS III, and one graph for all strings with two electrons in RAS III. In this case, the restrictions on combinations of strings become restrictions on combinations of graphs—a more efficient treatment computationally. Figure 4 displays string graphs for $n_\alpha = n_\beta = 3$ and $n = 6$ for at most 2 electrons in RAS III, orbitals 4-6. Graph (a) represents all walks with two electrons in RAS III; graph (b) gives all walks with one electron in RAS III; and graph (c) gives the one walk with no electrons in RAS III. If only two electrons are allowed in RAS III, it is clear that alpha strings of graph a can be combined only with the beta string from graph (c), and alpha strings of graph (b) can be combined with beta strings of graphs (b) and (c). An alpha string from graph (c) can be combined with beta strings from graphs (a), (b), and (c).

Figure 4: String graphs for $n_{\alpha/\beta} = 3, n = 6$ with at most two electrons in RAS III of orbitals 4-6. Vertex weights and arc weights are given for lexical ordering.
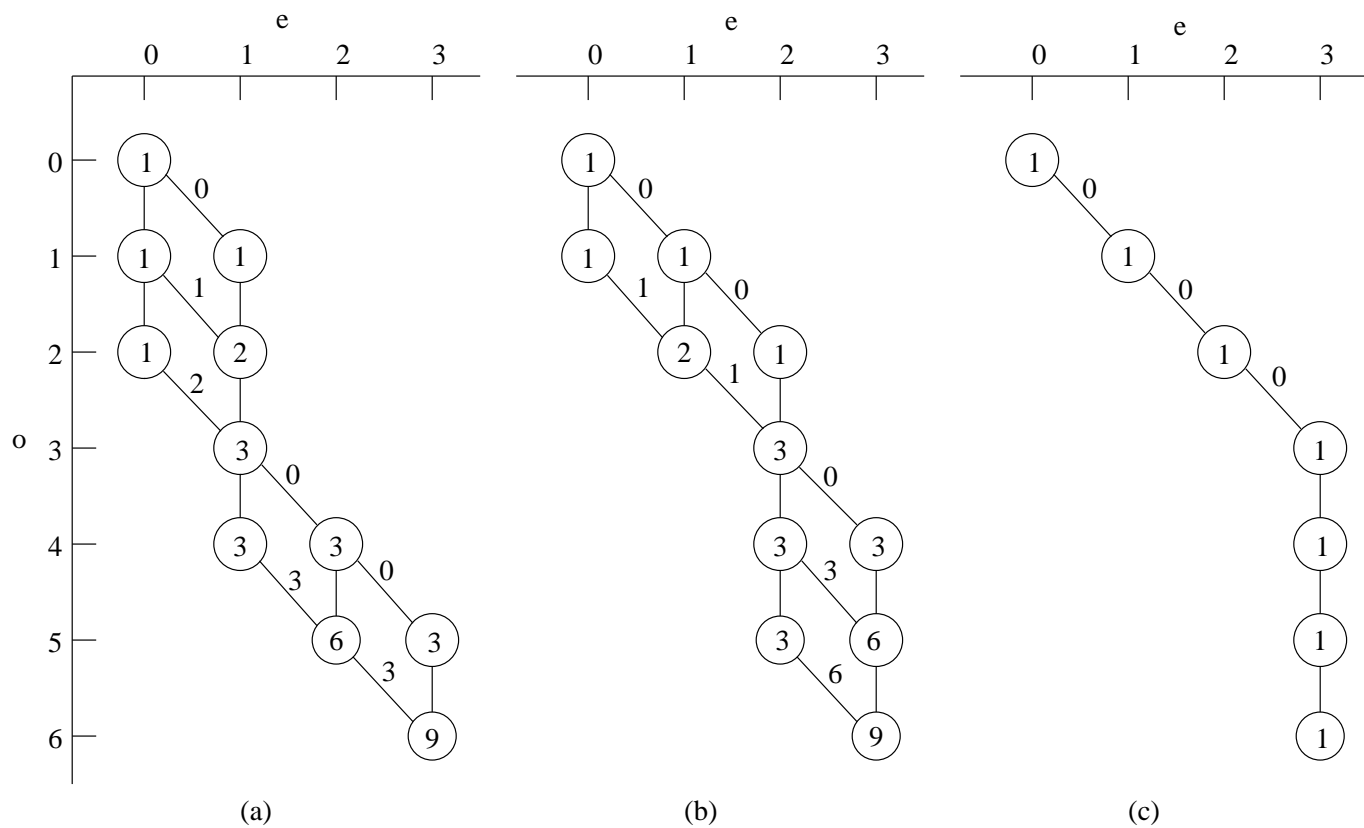
6.3.1   Olsen's Full CI σ Equations

We now turn our attention to the general formulation of the full CI problem in terms of alpha and beta strings. Later on, we will consider how to modify our results for RAS CI's. We begin by describing Olsen's expressions for σ, which is the action of the Hamiltonian $\hat{H}$ on the CI vector (or matrix) $C(I_\alpha, I_\beta)$. First we express $\hat{H}$ in second-quantized form (see section 5).

$$\hat{H} = \sum_{kl}^{n} h_{kl} \hat{E}_{kl} + \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij} \hat{E}_{kl} - \delta_{jk} \hat{E}_{il} \right) \tag{6.16}$$

where $\hat{E}_{kl}$ is the shift operator

$$\hat{E}_{kl} = \hat{E}_{kl}^{\alpha} + \hat{E}_{kl}^{\beta} \tag{6.17}$$

Again, σ is given by

$$\sigma(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta) \alpha(J_\alpha) | \hat{H} | \alpha(I_\alpha) \beta(I_\beta) \rangle C(J_\alpha, J_\beta) \tag{6.18}$$

As we will proceed to show, σ can be split up into three terms: one involving only beta components of the linear group generators (σ_1), one involving only alpha components of the generators (σ_2), and one involving mixtures of the two (σ_3). Inserting equation (6.16) into equation (6.18) yields

$$\sigma(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta)\alpha(J_\alpha)| \sum_{kl}^{n} h_{kl} \hat{E}_{kl} \tag{6.19}$$
$$+ \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il} \right) |\alpha(I_\alpha)\beta(I_\beta)\rangle C(J_\alpha, J_\beta)$$

Now expanding the shift operators according to equation (6.17), we write σ as a sum of three terms

$$\sigma(I_\alpha, I_\beta) = \sigma_1(I_\alpha, I_\beta) + \sigma_2(I_\alpha, I_\beta) + \sigma_3(I_\alpha, I_\beta) \tag{6.20}$$

where

$$\sigma_1(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta)\alpha(J_\alpha)| \sum_{kl}^{n} h_{kl} \hat{E}_{kl}^{\beta} \tag{6.21}$$
$$+ \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij}^{\beta}\hat{E}_{kl}^{\beta} - \delta_{jk}\hat{E}_{il}^{\beta} \right) |\alpha(I_\alpha)\beta(I_\beta)\rangle C(J_\alpha, J_\beta)$$

and

$$\sigma_2(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta)\alpha(J_\alpha)| \sum_{kl}^{n} h_{kl} \hat{E}_{kl}^{\alpha} \tag{6.22}$$
$$+ \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij}^{\alpha}\hat{E}_{kl}^{\alpha} - \delta_{jk}\hat{E}_{il}^{\alpha} \right) |\alpha(I_\alpha)\beta(I_\beta)\rangle C(J_\alpha, J_\beta)$$

and

$$\sigma_3(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \langle \beta(J_\beta)\alpha(J_\alpha)| \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij}^{\alpha}\hat{E}_{kl}^{\beta} + \hat{E}_{ij}^{\beta}\hat{E}_{kl}^{\alpha} \right) |\alpha(I_\alpha)\beta(I_\beta)\rangle C(J_\alpha, J_\beta) \tag{6.23}$$

Obviously, these expressions can be simplified further. First, we observe that the expression for σ_1 contains no α operators. Therefore, $\sigma_1(I_\alpha, I_\beta) = 0$ unless $J_\alpha = I_\alpha$. Using this fact, and integrating out the α part, we obtain

$$\sigma_1(I_\alpha, I_\beta) = \sum_{J_\beta} \langle \beta(J_\beta)| \sum_{kl}^{n} h_{kl} \hat{E}_{kl}^{\beta} + \frac{1}{2} \sum_{ijkl}^{n} (ij|kl) \left(\hat{E}_{ij}^{\beta}\hat{E}_{kl}^{\beta} - \delta_{jk}\hat{E}_{il}^{\beta} \right) |\beta(I_\beta)\rangle C(I_\alpha, J_\beta) \tag{6.24}$$

Taking out the Kronecker delta term and rearranging, we have

$$
\begin{aligned}
\sigma_1(I_\alpha, I_\beta) =\ & \sum_{J_\beta} \sum_{kl}^{n} h_{kl} \langle \beta(J_\beta) | \hat{E}_{kl}^{\beta} | \beta(I_\beta) \rangle C(I_\alpha, J_\beta) \qquad (6.25) \\
& - \frac{1}{2} \sum_{J_\beta} \sum_{ijl}^{n} (ij|jl) \langle \beta(J_\beta) | \hat{E}_{il}^{\beta} | \beta(I_\beta) \rangle C(I_\alpha, J_\beta) \\
& + \frac{1}{2} \sum_{J_\beta} \sum_{ijkl}^{n} (ij|kl) \langle \beta(J_\beta) | \hat{E}_{ij}^{\beta} \hat{E}_{kl}^{\beta} | \beta(I_\beta) \rangle C(I_\alpha, J_\beta)
\end{aligned}
$$

Now the second term can be combined with the first to give equation (9a) of reference [16].

$$
\begin{aligned}
\sigma_1(I_\alpha, I_\beta) =\ & \sum_{J_\beta} \sum_{kl}^{n} \langle \beta(J_\beta) | \hat{E}_{kl}^{\beta} | \beta(I_\beta) \rangle \left[h_{kl} - \frac{1}{2} \sum_{j}^{n} (kj|jl) \right] C(I_\alpha, J_\beta) \quad (6.26) \\
& + \frac{1}{2} \sum_{J_\beta} \sum_{ijkl}^{n} \langle \beta(J_\beta) | \hat{E}_{ij}^{\beta} \hat{E}_{kl}^{\beta} | \beta(I_\beta) \rangle (ij|kl) C(I_\alpha, J_\beta)
\end{aligned}
$$

Similarly, σ_2 can be simplified to

$$
\begin{aligned}
\sigma_2(I_\alpha, I_\beta) =\ & \sum_{J_\alpha} \sum_{kl}^{n} \langle \alpha(J_\alpha) | \hat{E}_{kl}^{\alpha} | \alpha(I_\alpha) \rangle \left[h_{kl} - \frac{1}{2} \sum_{j}^{n} (kj|jl) \right] C(J_\alpha, I_\beta) \quad (6.27) \\
& + \frac{1}{2} \sum_{J_\alpha} \sum_{ijkl}^{n} \langle \alpha(J_\alpha) | \hat{E}_{ij}^{\alpha} \hat{E}_{kl}^{\alpha} | \alpha(I_\alpha) \rangle (ij|kl) C(J_\alpha, I_\beta)
\end{aligned}
$$

For efficient implementation, it is convenient to precompute the quantities

$$
h'_{kl} = h_{kl} - \frac{1}{2} \sum_{j}^{n} (kj|jl) \qquad (6.28)
$$

Finally, we simplify σ_3. It may be rewritten as

$$
\begin{aligned}
\sigma_3(I_\alpha, I_\beta) =\ & \frac{1}{2} \sum_{J_\alpha, J_\beta} \sum_{ijkl}^{n} \langle \beta(J_\beta)\alpha(J_\alpha) | \hat{E}_{ij}^{\alpha} \hat{E}_{kl}^{\beta} | \alpha(I_\alpha)\beta(I_\beta) \rangle (ij|kl) C(J_\alpha, J_\beta) \\
& + \frac{1}{2} \sum_{J_\alpha, J_\beta} \sum_{ijkl}^{n} \langle \beta(J_\beta)\alpha(J_\alpha) | \hat{E}_{ij}^{\beta} \hat{E}_{kl}^{\alpha} | \alpha(I_\alpha)\beta(I_\beta) \rangle (ij|kl) C(J_\alpha, J_\beta)
\end{aligned}
$$

$$
(6.29)
$$

Since we sum over all $ijkl$, we can permute i and j with k and l. We can also swap $\hat{E}_{ij}^{\alpha}$ and $\hat{E}_{kl}^{\beta}$ as can easily be verified. This yields equation (9c) from reference [16].

$$\sigma_3(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \sum_{ijkl}^{n} \langle \beta(J_\beta)|\hat{E}_{ij}^{\beta}|\beta(I_\beta)\rangle \langle \alpha(J_\alpha)|\hat{E}_{kl}^{\alpha}|\alpha(I_\alpha)\rangle (ij|kl) C(J_\alpha, J_\beta)$$

(6.30)

Thus we have written the action of the Hamiltonian on the current CI vector in terms of alpha and beta strings and alpha and beta shift operators. The product σ is written as a sum of three terms: the first (σ_1) involves only beta shift operators, the second (σ_2) involves only alpha shift operators, and the third (σ_3) involves both alpha and beta shift operators. Note that, except for the factor $C(I_\alpha, J_\beta)$, σ_1 is independent of I_α so that the algorithm for computing σ_1 is vectorizable (see below). Analogous results hold for σ_2. This situation does not obtain in the computation of σ_3, however, which is the rate-limiting step.

6.4 Full CI Algorithm

It is easy to vectorize the formation of σ_1 since each element $\sigma_1(I_\alpha, I_\beta)$ can be written as independent of I_α apart from multiplication by a factor $C(I_\alpha, J_\beta)$. The vectorized algorithm for the evaluation of σ_1, adapted from reference [16], appears in Figure 5. An analogous algorithm can be used to obtain σ_2. However, we can also obtain σ_2 for $M_s = 0$ states by

$$\sigma_2(I_\alpha, I_\beta) = (-1)^S \sigma_1(I_\beta, I_\alpha)$$

(6.31)

as proven by Olsen and co-workers [16].

A fairly straightforward algorithm for evaluating σ_3 is presented in Figure 6. The vectorized algorithm for the evaluation of σ_3 is presented in Figure 7. This algorithm makes use of a gather and scatter operation to avoid indirect addressing.

For $M_s = 0$, an improvement to the σ_3 algorithm can be made by utilizing an identity similar to equation (6.31). The $ijkl$th component of σ_3 is related to the

Figure 5: Vectorized Algorithm for σ_1.

loop over I_β
      Set array $F(J_\beta) = 0$
      Loop over excitations $\hat{E}^\beta_{kl}$ from $|\beta(I_\beta)\rangle$
            $|\beta(K_\beta)\rangle = \mathrm{sgn}(kl)\hat{E}^\beta_{kl}|\beta(I_\beta)\rangle$
            $F(K_\beta) += \mathrm{sgn}(kl)h'_{kl}$
            Loop over excitations $\hat{E}^\beta_{ij}$ from $|\beta(K_\beta)\rangle$
                  $|\beta(J_\beta)\rangle = \mathrm{sgn}(ij)\hat{E}^\beta_{ij}|\beta(K_\beta)\rangle$
                  $F(J_\beta) += (1/2)\mathrm{sgn}(kl)\mathrm{sgn}(ij)(ij|kl)$
            end loop over $\hat{E}^\beta_{ij}$
      end loop over $\hat{E}^\beta_{kl}$
      $\sigma_1(I_\alpha, I_\beta) = \sum_{J_\beta} F(J_\beta)C(I_\alpha, J_\beta)$; vect'd over I_α
end loop over I_β

Figure 6: Simple Algorithm for σ_3.

loop over I_α
      loop over $|\alpha(J_\alpha)\rangle = \mathrm{sgn}(kl)\hat{E}^\alpha_{kl}|\alpha(I_\alpha)\rangle$
            loop over I_β
                  loop over $|\beta(J_\beta)\rangle = \mathrm{sgn}(ij)\hat{E}^\beta_{ij}|\beta(I_\beta)\rangle$
                        $\sigma_3(I_\alpha, I_\beta) += \mathrm{sgn}(ij)\mathrm{sgn}(kl)(ij|kl)C(J_\alpha, J_\beta)$

Figure 7: Vectorized Algorithm for σ_3.

loop over kl
      set up $L(I)$, $R(I)$, and sgn(I), defined by
      $|\alpha[L(I)]\rangle = \hat{E}_{kl}^{\alpha}|\alpha[R(I)]\rangle$sgn$(I)$
      $C'(I, J_\beta) = C[L(I), J_\beta]sgn(I)$; vectorized gather
      loop over I_β
            loop over excitations $\hat{E}_{ij}^{\beta}$ from $|\beta(I_\beta)\rangle$
                  $|\beta(J_\beta)\rangle = sgn(ij)\hat{E}_{ij}^{\beta}|\beta(I_\beta)\rangle$
                  $F(J_\beta)+ = sgn(ij)(ij|kl)$
            end loop over $\hat{E}_{ij}^{\beta}$
            $V(I) = \sum_{J_\beta} F(J_\beta)C'(I, J_\beta)$; vect'd over I
            $\sigma_3[R(I), I_\beta]+ = V(I)$; vectorized scatter
      end loop over I_β
end loop over kl

$klij$th component via

$$\sigma_3^{ijkl}(I_\alpha, I_\beta) = (-1)^S \sigma_1^{klij}(I_\beta, I_\alpha) \tag{6.32}$$

References

[1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York, 1989.

[2] W. Heisenberg, Z. Phys. **33**, 879 (1925).

[3] E. Schrödinger, Ann. d. Phys. **79**, 734 (1926).

[4] C. Eckart, Phys. Rev. **28**, 711 (1926).

[5] J. Paldus, J. Chem. Phys. **61**, 5321 (1974).

[6] R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **95**, 386 (1983).

[7] I. Shavitt, The method of configuration interaction, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 189–275, Plenum Press, New York, 1977.

[8] P. E. M. Siegbahn, The direct ci method, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson, pages 189–207, D. Reidel, Dordrecht, 1983.

[9] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[10] J. C. Slater, Phys. Rev. **38**, 1109 (1931).

[11] W. Ritz, J. Reine Angew. Math. **135**, 1 (1909).

[12] R. J. Bartlett, C. E. Dykstra, and J. Paldus, Coupled-cluster methods for molecular calculations, in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, edited by C. E. Dykstra, pages 127–159, D. Reidel, Dordrecht, 1984.

[13] L. Pauling and E. B. Wilson, Jr., *Introduction to Quantum Mechanics with Applications to Chemistry*, McGraw-Hill, New York, 1935.

[14] J. K. L. MacDonald, Phys. Rev. **43**, 830 (1933).

[15] E. A. Hylleraas and B. Undheim, Z. Phys. **65**, 759 (1930).

[16] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, J. Chem. Phys. **89**, 2185 (1988).

[17] J. T. Fermann, C. D. Sherrill, T. D. Crawford, and H. F. Schaefer, J. Chem. Phys. **100**, 8132 (1994).

[18] P. J. Knowles and N. C. Handy, J. Chem. Phys. **91**, 2396 (1989).

[19] J. Olsen, P. Jørgensen, and J. Simons, Chem. Phys. Lett. **169**, 463 (1990).

[20] A. O. Mitrushenkov, Chem. Phys. Lett. **217**, 559 (1994).

[21] R. P. Hosteny, T. H. Dunning, R. R. Gilman, A. Pipano, and I. Shavitt, J. Chem. Phys. **62**, 4764 (1975).

[22] V. R. Saunders and J. H. van Lenthe, Mol. Phys. **48**, 923 (1983).

[23] E. R. Davidson, Perspectives on *ab initio* calculations, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd, VCH Publishers, New York, 1990.

[24] W. Duch and G. H. F. Diercksen, J. Chem. Phys. **101**, 3018 (1994).

[25] S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).

[26] A. Povill, R. Caballol, J. Rubio, and J. P. Malrieu, Chem. Phys. Lett. **209**, 126 (1993).

[27] W. Duch, *GRMS or Graphical Representation of Model Spaces*, volume 1, Springer-Verlag, Berlin, 1986.

[28] G. E. Scuseria, C. L. Janssen, and H. F. Schaefer, J. Chem. Phys. **89**, 7382 (1988).

[29] N. C. Handy, Chem. Phys. Lett. **74**, 280 (1980).

Index